

**(12) INTERNATIONAL APPLICATION, PUBLISHED UNDER THE PATENT COOPERATION  
TREATY (PCT)**

**(19) World Intellectual Property Organization**  
International Bureau

[logo] [barcode]

**(43) International Publication Number**  
18 April 2002 (18.04.2002)

**PCT**      **(10) International Publication Number**  
**WO 02/31173 A2**

**(51) International Patent Classification<sup>7</sup>:** C12P  
17/06, C12N 15/11, 15/29, 15/52, 15/67, 15/74,  
15/79, 1/21, 5/04, 5/06, A23L 1/29, 1/302, A01H  
5/00, 5/08, 13/00, 15/00

**(21) International Application Number:**  
PCT/EP01/10779

**(22) International Filing Date:**  
18 September 2001 (18.09.2001)

**(25) Language in which the application was filed:**  
German

**(26) Language in which the application is  
published:** German

**(30) Priority data:**  
100 46 462.9 19 September 2000  
(19.09.2000) DE

**(71) Applicants (for all Designated States except for  
US):** SUNGENE GMBH & CO. KGaA [DE/DE];  
06468 Gatersleben (DE).

**(72) Inventor(s): and**

**(75) Inventor(s)/Applicant(s) (only for US):** GEIGER,  
Michael [DE/DE]; Neuer Weg 15, 06484  
Quedlinburg (DE). EBNETH, Marcus [DE/DE];  
Katzbachstr. 36, 10965 Berlin (DE). KUNZE, Irene  
[DE/DE]; Mühlenweg 11, 06466 Gatersleben  
(DE).

**(74) Agent:** DÖRPER, Thomas; Basf  
Aktiengesellschaft; 67056 Ludwigshafen (DE).

**(81) Designated States (national):** AE, AG, AL,  
AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA,  
CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ,  
EC, EE, ES, FI, GB, OD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,  
MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU,  
SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ,  
UA, UG, US, UZ, VN, YU, ZA, ZW.

**(84) Designated States (regional):** ARIPO Patent  
(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ,  
UG, ZW), European Patent (AM, AZ, BY, KG,  
KZ, MD, RU, TJ, TM), European Patent (AT,  
BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI Patent (BF,  
BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,  
MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be  
published again after receipt of the report

*For explanation of two-letter codes and other  
abbreviations, see explanations ("Guidance Notes on  
Codes and Abbreviations") at the beginning of each  
regular issue of the PCT Gazette.*

[barcode]

**(54) Title [English in original]:** IMPROVED METHOD FOR THE BIOSYNTHESIS OF VITAMIN E

**(54) Title [translation of German]:** IMPROVED METHODS FOR VITAMIN E BIOSYNTHESIS

**(57) Abstract [provided in English in original]:** The invention relates to improved methods for the biosynthesis of vitamin E. Said methods are characterized by the inhibition of the catabolization of homogentisate to maleylacetoacetate and fumarylacetoacetate and then to fumarate and acetoacetate. The invention also relates to the combination of this inhibition with methods that increase the provision of homogentisate or that promote the conversion of homogentisate to vitamin E. The invention further relates to nucleic acid constructs and vectors, which can be used to implement the inventive methods, in addition to transgenic plant organisms produced from said constructs and vectors.

**(57) Abstract [translation of German]:** The invention relates to improved methods for biosynthesis of Vitamin E. These methods are characterized by inhibition of homogentisate degradation via maleylacetoacetate and fumarylacetoacetate to form fumarate and acetoacetate. The invention also relates to the combination of this inhibition with methods that increase the availability of homogentisate or promote conversion of homogentisate to Vitamin E. The invention relates to nucleic acid constructs and vectors which can be used to implement the methods according to the invention, as well as transgenic plant organisms which are produced therefrom.

WO 02/31173 A2

## DECLARATION

I, Cathy Flick, technical translator for German-American Business Translation, do hereby declare that I am conversant with the English and German languages and am a competent translator thereof. I declare further that to the best of my knowledge and belief the following translation of International Publication Number WO 02 /31173 A2 is a true and correct translation made by me of the document in the German language attached hereto.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Signed this 16th day of January, 2003.

Cathy Flick

## Improved methods for Vitamin E biosynthesis

## Specification

The invention relates to improved methods for biosynthesis of Vitamin E. These methods are characterized by inhibition of homogentisate (HG) degradation via maleylacetoacetate (MAA) and fumarylacetoacetate (FAA) to form fumarate and acetoacetate. The invention also relates to the combination of this inhibition with methods that further increase the availability of homogentisate or promote conversion of homogentisate to Vitamin E.

Homogentisate is an important metabolite. It is a degradation product of the amino acids tyrosine and phenylalanine. In humans and animals, homogentisate is further broken down to maleylacetoacetate, then to fumarylacetoacetate and then to fumarate and acetoacetate. Plants and other photosynthetic microorganisms also use homogentisate as a starting product for synthesis of tocopherols and tocotrienols.

The eight naturally occurring compounds with Vitamin E activity are derivatives of 6-chromanol (Ullmann's Encyclopedia of Industrial Chemistry, Vol. A 27 (1996), VCH Verlagsgesellschaft, Chapter 4, 478-488, Vitamin E). Members of the tocopherol group (1a-d) have a saturated side chain, while members of the tocotrienol group (2a-d) have an unsaturated side chain:

[display formula on German p. 1] (1)

1a,  $\alpha$ -tocopherol:  $R^1 = R^2 = R^3 = CH_3$

1b,  $\beta$ -tocopherol [148-03-8]:  $R^1 = R^3 = CH_3$ ,  $R^2 = H$

1c,  $\gamma$ -tocopherol [54-28-4]:  $R^1 = H$ ,  $R^2 = R^3 = CH_3$

1d,  $\delta$ -tocopherol [119-13-1]:  $R^1 = R^2 = H$ ,  $R^3 = CH_3$

[display formula on German p. 1] (2)

2a,  $\alpha$ -tocotrienol [1721-51-3]:  $R^1 = R^2 = R^3 = CH_3$

2b,  $\beta$ -tocotrienol [490-23-3]:  $R^1 = R^3 = CH_3$ ,  $R^2 = H$

2c,  $\gamma$ -tocotrienol [14101-61-2]:  $R^1 = H$ ,  $R^2 = R^3 = CH_3$

2d,  $\delta$ -tocotrienol [25612-59-3]:  $R^1 = R^2 = H$ ,  $R^3 = CH_3$

In the present invention, "Vitamin E" means all eight above-mentioned tocopherols and tocotrienols with Vitamin E activity.

These compounds with Vitamin E activity are important natural fat-soluble antioxidants. Vitamin E deficiency results in pathophysiological conditions in humans and animals. Epidemiological studies have shown that a nutritional supplement with Vitamin E reduces the risk of cardiovascular diseases or cancer. A positive effect on the immune system and prevention of general age-related degeneration have been described (Traber M.G., Sies H.; Ann Rev Nutr. 1996;16:321-47). In these cases, the function of Vitamin E is probably to stabilize the biological membranes as well as to reduce free radicals as occur, for example, in lipid oxidation of polyunsaturated fatty acids (PUFA).

The function of Vitamin E in the plants themselves has been hardly studied. However, it seems possible that they have an important function in the stress response of the plant, especially response to oxidative stress. Elevated Vitamin E levels have been connected with high stability and shelf life of plant products. Adding Vitamin E to animal feed has a positive effect on the quality of the meat and the shelf life of meat and meat products for pork, beef, and poultry, for example.

Vitamin E compounds thus have high economic value as additives in the food and feed sector, in pharmaceutical formulations, and in cosmetic applications.



In nature, Vitamin E is synthesized exclusively by plants and other photosynthetically active organisms (for example, cyanobacteria). The Vitamin E content varies considerably. Most basic food plants (for example, wheat, rice, corn, potato) have only very small amounts of Vitamin E (Hess, Vitamin E,  $\alpha$ -Tocopherol, in *Antioxidants in Higher Plants*, R. Ascher and J. Hess, editors, 1993, CRC Press, Boca Raton, pp. 111-134). Oilseeds generally have a distinctly higher Vitamin E content, where  $\beta$ -,  $\gamma$ -,  $\delta$ -tocopherols predominate. The recommended daily dosage of Vitamin E is 15-30 mg.

Fig. 1 shows a biosynthesis pathway for tocopherols and tocotrienols.

During biosynthesis, homogentisic acid (homogentisate; HG) binds to phytyl pyrophosphate (PPP) or geranylgeranyl pyrophosphate to form precursors of  $\alpha$ -tocopherol and  $\alpha$ -tocotrienol: 2-methyl phytylhydroquinone or 2-methyl geranylgeranyl hydroquinone. In the methylation step, with S-adenosylmethionine as the methyl group donor, first 2,3-dimethyl-6-phytylhydroquinone is formed, then cyclization results in  $\gamma$ -tocopherol, and  $\alpha$ -tocopherol is formed by further methylation. The  $\beta$ - and  $\delta$ -tocopherol can also be synthesized from 2-methyl-phytylhydroquinone by methylation.

Not much is currently known about increasing the metabolic flux in order to increase the tocopherol or tocotrienol content in transgenic organisms, for example, in transgenic plants by overexpression of individual biosynthesis genes.

WO 97/27285 describes modification of the tocopherol content by enhanced expression or by downregulation of the enzyme *p*-hydroxyphenyl pyruvate dioxygenase (HPPD).

WO 99/04622 describes gene sequences coding for a  $\gamma$ -tocopherol methyltransferase from *Synechocystis* PCC6803 and *Arabidopsis thaliana* and their incorporation into transgenic plants.

WO 99/23231 shows that expression of a geranylgeranyl oxidoreductase in transgenic plants results in increased tocopherol biosynthesis.

WO 00/10380 shows a change in the Vitamin E composition when 2-methyl-6-phytylplastoquinol methyltransferase is used.

Shintani and DellaPenna have shown that overexpression of  $\gamma$ -tocopherol methyltransferase can distinctly raise the Vitamin E content (Shintani and DellaPenna, Science 282 (5396): 2098-2100, 1998).

All reactions for Vitamin E biosynthesis go through the homogentisate. Previous studies have been limited mainly to overexpression of genes for Vitamin E or homogentisate biosynthesis (see above). Little attention has been paid so far to the competing reactions that break down the homogentisate and thus detract from Vitamin E biosynthesis.

Degradation of homogentisate through maleylacetoacetate and fumarylacetoacetate to form fumarate and acetoacetate has been described for organisms that are not photosynthetically active, primarily animal organisms (Fernandez-Canon J.M. et al., Proc Natl Acad Sci USA, 1995; 92 (20): 9132-9136). Animal organisms use this metabolic pathway to break down aromatic amino acids, which are mainly taken up via food. Its function and relevance in plants, however, is unclear. The reactions are catalyzed by homogentisate 1,2-dioxygenase (HGD; EC: 1.13.11.5), maleylacetoacetate isomerase (MAAI; EC: 5.2.1.2.) and fumarylacetoacetate hydrolase (FAAH; EC: 3.7.1.2).

The gene for homogentisate 1,2-dioxygenase (HGD) from *Arabidopsis thaliana* is known (GenBank Accession Number AF130845). The gene for fumarylacetoacetate hydrolase from *Arabidopsis thaliana*, based on homology to fumarylacetoacetate hydrolase from *Emmericella nidulans* (gb|L41670), has already been annotated as similar to the latter (GenBank Accession Number AC002131). However, in the corresponding GenBank entry it is explicitly pointed out that the annotation is based solely on similarity and not on experimental data. The gene for maleylacetoacetate isomerase (MAAI) from *Arabidopsis* was available as a gene in GenBank (AC005312), but annotated as a putative glutathione S-transferase. An MAAI was known from *Emmericella nidulans* (GenBank Accession Number EN 1837).

In an abstract (Abstract No. 413) at the 1999 annual meeting of the American Society of Plant Physiologists (24-28 July 1999, Baltimore, USA), Tsegaye et al. speculate about an advantage in combining crossing of HPPD-overexpressing plants with plants in which HGD is downregulated by means of an antisense approach.

Despite some success, optimization of Vitamin E biosynthesis is still needed. The aim of the invention is to provide further methods to affect the Vitamin E biosynthesis pathway and to thus result in further beneficial transgenic plants with elevated Vitamin E content.

The aim has been achieved by identification of the homogentisate — maleylacetoacetate — fumarylacetoacetate fumarate degradation pathway as an important pathway competing with the Vitamin E biosynthesis pathway. It was found that inhibition of this degradation pathway leads to optimization of Vitamin E biosynthesis.

A first subject matter of the invention therefore relates to methods for Vitamin E production by decreasing HGD, MAAI, FAAH activity. A combination of the described inhibition of the homogentisate degradation pathway with other methods that lead to improved Vitamin E biosynthesis, by promoting conversion of homogentisate to Vitamin E, has proven to be particularly advantageous. This may be realized because of increased availability of the reaction partners, or because of the resulting increased conversion of homogentisate. For example, this effect can be achieved with overexpression of homogentisate phytyltransferase (HGPT), geranylgeranyl oxidoreductase, 2-methyl-6-phytylplastoquinol methyltransferase, or  $\gamma$ -tocopherol methyltransferase.

A combination with genes that promote homogentisate formation is also advantageous, such as for example the HPPD or TyrA gene.

The degradation pathway for homogentisate via maleylacetoacetate and fumarylacetoacetate to form fumarate and acetoacetate can be inhibited in many different ways.

A subject matter of the invention relates to nucleic acid constructs that contain at least one nucleic acid sequence (anti-MAAI/FAAH) capable of inhibiting the maleylacetoacetate — fumarylacetoacetate fumarate pathway, or a functional equivalent thereof.

A further subject matter of the invention relates to the above-described nucleic acid constructs that contain, besides the anti-MAAI/FAAH nucleic acid sequence, in addition at least one nucleic acid sequence (pro-HG) capable of increasing biosynthesis of homogentisate (HG), or a functional equivalent thereof, or at least one nucleic acid sequence (provitamin E)

capable of increasing Vitamin E biosynthesis starting from homogentisate, or a functional equivalent thereof, or a combination of pro-HG and provitamin E or their functional equivalents.

A further subject matter of the invention relates to nucleic acid constructs that contain a nucleic acid sequence (anti-HGD) capable of inhibiting homogentisate 1,2-dioxygenase, or a functional equivalent thereof.

The invention also relates to the aforementioned anti-HGD nucleic acid constructs, which contain, besides the anti-HGD nucleic acid sequence, additionally at least one nucleic acid sequence [*sic - "coding" is missing word*] for a bifunctional chorismate mutase — prephenate dehydrogenase (TyrA) or functional equivalents thereof, or at least one nucleic acid sequence (provitamin E) capable of increasing Vitamin E biosynthesis starting from homogentisate, or functional equivalents thereof, or a combination of provitamin E and TyrA sequences or functional equivalents thereof.

TyrA codes for a bifunctional chorismate mutase — prephenate dehydrogenase from *E. coli*, a hydroxyphenylpyruvate synthase that includes the enzyme activities of a chorismate mutase and prephenate dehydrogenase and converts chorismate to hydroxyphenyl pyruvate, the starting material for homogentisate (Christendat D., Turnbull J.L. Biochemistry. 1999 Apr 13;38(15): 4782-93; Christopherson R.I., Heyde E., Morrison J.F. Biochemistry. 1983 Mar 29;22(7): 1650-6.).

The invention also relates to nucleic acid constructs that contain at least one nucleic acid sequence (pro-HG) capable of increasing biosynthesis of homogentisate (HG), or a functional equivalent thereof, and at least one nucleic acid sequence (provitamin E) capable of increasing Vitamin E biosynthesis starting from homogentisate, or a functional equivalent thereof.

The invention in addition relates to functional analogs of the above-mentioned nucleic acid constructs. Here functional analogs mean, for example, a combination of individual nucleic acid sequences

1. on one polynucleotide (multiple constructs)
2. on several polynucleotides in one cell (cotransformation)

3. by crossing different transgenic plants that each contain at least one of the aforementioned nucleotide sequences.

The nucleic acid sequences contained in the nucleic acid constructs are preferably operably linked with genetic control sequences.

The transformation of plants with a pro-HG coding construct according to the invention leads to an increase in homogentisate formation. Simultaneous transformation with anti-HGD or anti-MAAI/FAAH, in particular the anti-MAAI construct, avoids an undesirable loss of these metabolites. So an increased amount of homogentisate is available in the transgenic plants to form Vitamin E, for example, to form tocopherols through the intermediates methyl-6-phytylquinol and 2,3-dimethyl phytylquinol (see Fig. 1). Both pro-HG and anti-MAAI/FAAH or anti-HGD lead to increased availability of homogentisate for Vitamin E biosynthesis. Conversion of homogentisate to Vitamin E can be improved by combined transformation with a construct coding for provitamin E, and in addition increases Vitamin E biosynthesis.

An "increase" in homogentisate biosynthesis in this context is to be broadly interpreted, and includes the increase in the homogentisate (HG) biosynthesis activity in the plant transformed with a pro-HG construct according to the invention, or in part of the plant or plant tissue. The invention includes various strategies for raising the HG biosynthesis activity. The person skilled in the art knows that a number of different methods are available for influencing HG biosynthesis activity in a desired way. The methods described below in this respect are by way of example and are not to be considered restrictive.

The preferred strategy according to the invention includes the use of a nucleic acid sequence (pro-HG) that can be transcribed and translated to a polypeptide, which increases HG biosynthesis activity. Examples of such nucleic acid sequences are *p*-hydroxyphenylpyruvate dioxygenase (HPPD) from various organisms, or the bacterial TyrA gene product. Besides the described artificial expression of known genes, their activity can also be increased by mutagenesis of the polypeptide sequence. Increased transcription and translation of endogenous genes can also be achieved by using artificial transcription factors of the zinc finger protein type, for example (Beerli RR et al., Proc Natl Acad Sci USA 2000; 97 (4): 1495-500). These factors are taken up by the regulatory regions of the endogenous genes and, depending on the factor design,

cause expression or repression of the endogenous gene.

Particularly preferred for pro-HG is the use of [*"durch"= "by" in German seems to be a typo, compare with last sentence of next paragraph*] nucleic acids that code for polypeptides as specified by SEQ ID NO: 8, 11, or 16, particularly preferred are nucleic acids with sequences described by SEQ ID NO: 7, 10, or 15.

"Increase" in Vitamin E biosynthesis activity should be analogously understood, where here genes are used with activity that promotes conversion of homogentisate to Vitamin E (tocopherols, tocotrienols) or synthesis of reaction partners of homogentisate, such as for example phytyl pyrophosphate or geranylgeranyl pyrophosphate. Examples can include homogentisate phytyltransferase (HGPT), geranylgeranyl oxidoreductase, 2-methyl-6-phytylplastoquinol methyltransferase, and  $\gamma$ -tocopherol methyltransferase. Particularly preferred is the use of nucleic acids that code for polypeptides as specified by SEQ ID NO: 14, 20, 22, or 24, particularly preferred are those with sequences described by SEQ ID NO: 13, 19, 21 or 23.

"Inhibition" should be broadly interpreted in the context with anti-MAAI/FAAH or anti-HGD and includes, depending on the different cell biological mechanisms, partial or essentially complete prevention or blocking of MAAI/FAAH or HGD enzyme activity in the plant transformed with an anti-MAAI/FAAH or anti-HGD construct according to the invention, or in part of the plant or plant tissue. Inhibition according to the invention also encompasses a quantitative decrease in active HGD, MAAI, or FAAH in the plant, down to essentially complete absence of HGD, MAAI, or FAAH protein (i.e., no detectable HGD or MAAI or FAAH enzyme activity, or no immunologically detectable HGD, MAAI, or FAAH).

The invention includes various strategies for decreasing or inhibiting HGD or MAAI or FAAH activity. The person skilled in the art knows that a number of different methods are available for influencing HGD or MAAI or FAAH gene expression or enzyme activity in a desired way.

The preferred strategy according to the invention is to use a nucleic acid sequence (anti-MAAI/FAAH or anti-HGD) which can be transcribed to an antisense nucleic acid sequence capable of inhibiting HGD or MAAI/FAAH activity.

for example, by inhibiting expression of endogenous HGD or MAAI or FAAH.

The anti-HGD or anti-MAAI/FAAH nucleic acid sequences according to the invention can, according to a preferred embodiment, contain HGD (anti-HGD) or MAAI or FAAH (anti-MAAI/FAAH) coding nucleic acid sequences inserted in the antisense orientation, or functionally equivalent fragments of the respective sequences.

Particularly preferred anti-HGD nucleic acid sequences include nucleic acid sequences which code for polypeptides containing an amino acid sequence as specified by SEQ ID NO: 3 or functional equivalents thereof. Particularly preferred are nucleic acid sequences specified by SEQ ID NO: 1, 2, or 12 or functional equivalents thereof.

Particularly preferred anti-MAAI/FAAH nucleic acid sequences include nucleic acid sequences which code for polypeptides containing an amino acid sequence as specified by SEQ ID NO: 5 and 18 or functional equivalents thereof. Particularly preferred are nucleic acid sequences as specified by SEQ ID NO: 4, 6, 9 or 17 or functional equivalents thereof, more particularly preferred are those with partial sequences shown with SEQ ID NO: 41 or 42 or functional equivalents thereof.

A preferred embodiment of the nucleic acid sequences according to the invention includes an HGD, MAAI, or FAAH sequence motif as specified by SEQ ID NO: 1, 2, 4, 6, 9, 12, 17, 41, or 42 in antisense orientation. This leads to increased transcription of nucleic acid sequences in the transgenic plant which are complementary to endogenous sequences coding for HGD, MAAI, or FAAH or a portion thereof, and which hybridize with the latter at the DNA or RNA level.

The antisense strategy can be advantageously combined with a ribozyme method. Ribozymes are catalytically active RNA sequences that bind to the antisense sequences, that catalytically cleave the target sequences (Tanner NK. FEMS Microbiol Rev. 1999; 23 (3): 257-75). This can increase the efficiency of an antisense strategy.

Further methods for inhibition of HGD or MAAI/FAAH expression include overexpression of homologous HGD or MAAI/FAAH nucleic acid sequences leading to cosuppression (Jorgensen et al., Plant Mol. Biol. 1996, 31 (5): 957-973), induction of specific RNA degradation by the plant using a viral

expression system (amplicon) (Angell, SM et al. , Plant J. 1999, 20(3): 357-362). These methods are also called "post-transcriptional gene silencing" (PTGS).

Further methods are introduction of nonsense mutations into the endogenous gene by introducing RNA/DNA oligonucleotides into the plant (Zhu et al., Nat. Biotechnol. 2000, 18(5): 555-558) or generation of knockout mutants using, for example, T-DNA-mutagenesis (Koncz et al., Plant Mol. Biol. 1992, 20(5):963-976) or homologous recombination (Hohn, B. and Puchta H, Proc. Natl. Acad. Sci. USA. 1999, 96: 8321-8323). A further method is gene overexpression or repression, also with specific DNA-binding factors, for example with the above-mentioned factors of the zinc finger transcription factor type. Factors in a cell that inhibit the target protein itself can also be used. The protein-binding factors can be, for example, aptamers (Famulok M and Mayer G. Curr Top Microbiol Immunol. 1999; 243: 123- 36).

Reference is hereby explicitly made to the papers described above and the methods for regulation of plant gene expression disclosed therein.

An anti-HGD or anti-MAAI/FAAH sequence according to the present invention is thus in particular selected from among:

- a) antisense nucleic acid sequences;
- b) antisense nucleic acid sequences combined with a ribozyme method
- c) nucleic acid sequences coding for homologous HGD or MAAI/FAAH and leading to cosuppression
- d) viral nucleic acid sequences and expression constructs causing HGD or MAAI/FAAH RNA degradation;
- e) nonsense mutants of endogenous nucleic acid sequences coding for HGD or MAAI/FAAH;
- f) nucleic acid sequences coding for knockout mutants;
- g) nucleic acid sequences suitable for homologous recombination;



- h) nucleic acid sequences coding for specific DNA- or protein-binding factors with anti-HGD or anti-MAAI/FAAH activity

wherein expression of each of these individual anti-HGD or anti-MAAI/FAAH sequences can cause "inhibition" of HGD or MAAI/FAAH activity according to the invention. Combined application of such sequences is also conceivable.

Nucleic acid construct or nucleic acid sequence according to the invention means, for example, a genomic or complementary DNA sequence or an RNA sequence as well as semisynthetic or completely synthetic analogs thereof. These sequences can be in linear or circular form, extrachromosomal or integrated into the genome. The pro-HG, provitamin E, anti-HGD or anti-MAAI/FAAH nucleotide sequences of the constructs according to the invention can be synthetically prepared or naturally extracted or can contain a mixture of synthetic and natural DNA components, and can consist of various heterologous HGD, MAAI/FAAH, pro-HG, or provitamin E gene segments from different organisms. The anti-HGD or anti-MAAI/FAAH sequence can be derived from one or several exons or introns, in particular exons of HGD, MAAI, or FAAH genes.

Artificial nucleic acid sequences are also suitable as long as they, as described above, impart the desired property, for example the increase in Vitamin E content in the plant by means of overexpression of at least one pro-HG and/or provitamin E gene and/or expression of an anti-HGD or MAAI/FAAH sequence in cultivated plants. For example, synthetic nucleotide sequences with codons can be produced that are preferred by the plants to be transformed. These codons preferred by the plants can be identified in the conventional way, based on codon usage of codons with the highest protein frequency. Such artificial nucleotide sequences can, for example, be determined by back-translation of proteins constructed using molecular modelling that exhibit HGD or MAAI/FAAH or proHG activity or provitamin E activity, or by *in vitro* selection. Particularly suitable are coding nucleotide sequences that were obtained by back-translation of a polypeptide sequence according to the codon usage specific for the host plant. Undesirable plant regulation mechanisms can be avoided, for example, by starting from the amino acid sequence of a bacterial pro-HG, for example the bacterial TyrA gene, and taking into account the plant codon usage, back-translating the DNA fragments, and from that the complete exogenous pro-HG sequence, optimized for use in the plant, can be prepared.

From that, a pro-HG enzyme is expressed which is either not amenable to plant regulation or is only slightly so, and as a result overexpression of enzyme activity can be achieved to its fullest potential.

All the nucleotide sequences mentioned above can be prepared in a known manner by chemical synthesis from the nucleotide components such as, for example, by fragment condensation of individual overlapping complementary nucleic acid component of the double helix. Chemical synthesis of oligonucleotides can, for example, be carried out in a known way by the phosphoramidite method (Voet and Voet, 2nd Edition, Wiley Press. New York, pp. 896-897). In preparation of a nucleic acid construct, various DNA fragments can be manipulated so that a nucleotide sequence with the correct reading direction and the correct reading frame is obtained. Adaptors or linkers can be attached to the fragments to link the nucleic acid fragments with each other. The attachment of synthetic oligonucleotides and filling in gaps using Klenow fragments of DNA polymerase and ligation reactions as well as general cloning methods are described in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press.

Functional equivalents of pro-HG or provitamin E sequences are sequences that, despite differing nucleotide sequences, still code for a protein with the desired functions according to the invention, i.e., for an enzyme with activity directly or indirectly increasing homogentisate formation (pro-HG), or for an enzyme with activity directly or indirectly promoting conversion of homogentisate to Vitamin E (provitamin E).

Functional equivalents of anti-HGD or anti-MAAI/FAAH include nucleotide sequences such that the HGD or MAAI/FAAH enzyme function is prevented in the transgenic plant to a sufficient extent. This can occur, for example, by hindering or preventing HGD or MAAI/FAAH processing, transport of HGD or MAAI/FAAH or their mRNA, inhibition of ribosome attachment, inhibition of RNA splicing, induction of an RNA-degrading enzyme and/or inhibition of translation elongation or termination. Direct repression of endogenous genes by DNA-binding factors, for example of the zinc finger transcription factor type, is also possible. Direct inhibition of the corresponding polypeptides, for example, by aptamers, is also feasible. Various examples are given above.

Functional equivalent means in particular also natural or artificial mutations of an originally isolated sequence coding for HGD or MAAI/FAAH or pro-HG or provitamin E which still exhibits the desired function. Mutations include substitutions, additions, deletions, exchanges, or insertions of one or more nucleotide residues. Thus for example, the present invention also encompasses nucleotide sequences that can be obtained by modification of the HGD or MAAI/FAAH or pro-HG or provitamin E nucleotide sequence. The aim of such modification can be, for example, further delimiting of the coding sequences contained therein or, for example, also insertion of more restriction enzyme cleavage sites or removal of superfluous DNA.

Where insertions, deletions, or substitutions are involved, such as for example transitions and transversions, techniques known per se such as *in vitro* mutagenesis, primer repair, restriction or ligation can be used. By means of manipulations such as for example restriction, "chewing-back" or filling in overhangs for "blunt ends," complementary ends of the fragments can be used for ligation.

Substitution means exchange of one or more amino acids for one or more amino acids. "Conservative" exchanges are preferably carried out, in which the substituted amino acid has properties similar to the properties of the original amino acid, for example exchange of Glu for Asp, Gln for Asn, Val for Ile, Leu for Ile, Ser for Thr.

Deletion is replacement of an amino acid by a direct bond. Preferred positions for deletions are the ends of polypeptides and the linkages between individual protein domains.

Insertions are entries of amino acids into the polypeptide chain, where formally a direct bond is replaced by one or more amino acids.

Homology between two proteins means that the amino acids are identical over the respective total protein lengths, which is calculated by comparison using the program algorithm GAP (UWGCG, University of Wisconsin, Genetic Computer Group), setting the following parameters:

Gap Weight: 12                      Length Weight: 4  
Average Match: 2.912              Average Mismatch: -2.003

A sequence that has at least 20% homology to a nucleic acid basis with the sequence SEQ ID NO. 6 means a sequence such that, when its sequence is compared with the sequence SEQ ID NO. 6 by the above program algorithm with the above parameter settings, homology of at least 20% results.

Functional equivalents derived from one of the nucleic acid sequences used in nucleic acid constructs or vectors according to the invention, for example by substitution, insertion, or deletion of amino acids or nucleotides, have at least 20% homology, preferably 40%, preferably at least 60%, preferably at least 80%, particularly preferably at least 90%.

Further examples of nucleic acid sequences used in nucleic acid constructs or vectors according to the invention can be easily found, for example, from among the different organisms for which the genomic sequence is known, such as for example *Arabidopsis thaliana*, by a homology comparison of the amino acid sequences or the corresponding back-translated nucleic acid sequences from data banks.

Functional equivalents also include variants such that their function is attenuated or enhanced compared with the starting gene or gene fragment, that is, for example pro-HG or provitamin E genes that code for a polypeptide variant with lower or higher enzymatic activity compared with the original gene.

Further suitable functionally equivalent nucleic acid sequences include sequences that code for fusion proteins, where for example a pro-HG or provitamin E polypeptide or a functionally equivalent portion thereof is a component of the fusion protein. The second portion of the fusion protein can, for example, be another polypeptide with enzyme activity (for example, another pro-HG or provitamin E polypeptide or a functionally equivalent portion thereof) or an antigene polypeptide sequence that can be used to detect pro-HG or provitamin E expression (for example, myc-tag or his-tag). However, then it preferably involves a regulatory protein sequence, such as for example a signal or transit peptide that directs the pro-HG or provitamin E protein to the desired site of action.

A further subject matter of the invention relates to recombinant vectors that include at least one nucleic acid construct according to the above definition, one nucleic acid sequence that codes for HGD, MAAI, or FAAH, or combinations of these options.

The nucleic acid sequences or nucleic acid constructs contained in the vectors are preferably operably linked with genetic control sequences.

Examples of vectors according to the invention can include the following types of expression constructs:

- a) 5'-plant-specific promoter / anti-HGD / terminator-3'
- b) 5'-plant-specific promoter / anti-MAAI/FAAH / terminator-3'
- c) 5'-plant-specific promoter / pro-HG / terminator-3'
- d) 5'-plant-specific promoter / provitamin E / terminator-3'

The invention explicitly also relates to vectors that can express polypeptides with HGD, MAAI, or FAAH activity. The sequences coding for these genes preferably come from plants, cyanobacteria, mosses, fungi, or algae. Sequences coding for polypeptides specified by SEQ ID NO 3, 5, and 18 are particularly preferred.

Here a pro-HG or provitamin E coding sequence, as well as sequences for expression of polypeptides with HGD, MAAI, or FAAH activity, can also be replaced by a sequence coding for a fusion protein made up from a transit peptide and the corresponding sequence.

Preferred examples include vectors and can contain one of the following expression constructs:

- a) 5'-35S-promoter / anti-MAAI/FAAH / OCS-terminator-3'
- b) 5'-35S-promoter / anti-HGD / OCS-terminator-3';
- c) 5'-legumin B promoter / pro-HG / NOS-terminator-3'
- d) 5'-legumin B promoter / provitamin E / NOS-terminator-3'

- e) 5'-legumin B promoter / HGD / NOS-terminator-3'
- f) 5'-legumin B promoter / MAAI / NOS-terminator-3'
- g) 5'-legumin B promoter / FAAH / NOS-terminator-3'

Here the pro-HG or provitamin E coding sequence can also be replaced by a sequence coding for a fusion protein made up from a transit peptide and pro-HG or provitamin E.

A cotransformation with more than one of the above indicated examples a) to g) may be required for the advantageous methods according to the invention for optimization of vitamin E biosynthesis. The transformation can also be advantageous with one or more vectors that each contain one combination of the above-indicated constructs. Preferred examples include vectors containing the following constructs:

- a) 5'-35S-promoter/ anti-MAAI/FAAH / OCS-terminator / legumin B promoter / pro-HG / NOS-terminator-3';
- b) 5'-35S-promoter/ anti-MAAI/FAAH / OCS-terminator / legumin B promoter / provitamin E / NOS-terminator-3';
- c) 5'-35S-promoter/ anti-HGD / OCS-terminator / legumin B promoter / provitamin E / NOS-terminator-3';
- d) 5'-35S-promoter/ pro-HG / OCS-Terminator / legumin B promoter / provitamin E / NOS-terminator-3';

Constructs a) to d) permit simultaneous transformation of the plant with pro-HG or provitamin E and anti-HGD or anti-MAAI/FAAH.

If the above-indicated recombination and cloning techniques are used, the nucleic acid constructs can be cloned into suitable vectors that enable them to replicate, for example in *E. coli*. Suitable cloning vectors are, among others, pBR332, pUC series, M13mp series, and pACYC184. Binary vectors that can replicate both in *E. coli* and in *Agrobacteria* are particularly suitable.

The nucleic acid constructs according to the invention are preferably inserted into suitable transformation vectors. Suitable vectors are, among others, described in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Chapters 6/7, pp. 71-119 (1993).

They are preferably cloned into a vector such as, for example, pBin19, pBinAR, pPZP200, or pPTV, that is suitable for transforming *Agrobacterium tumefaciens*. The *Agrobacteria* transformed with such a vector can then be used in a known way for transformation of plants, in particular cultivated plants such as for example rape, for example by bathing wounded leaves or leaf fragments in an *Agrobacteria* solution and then cultivating them in suitable media. Transformation of plants via *Agrobacteria* is known, among others, from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38. Transgenic plants integrating the nucleic acid constructs described above can be regenerated in a known way from the transformed cells of the wounded leaves or leaf fragments.

The nucleic acid sequences contained in the nucleic acid constructs and vectors according to the invention can be operably linked with at least one genetic control sequence. Genetic control sequences ensure, for example, transcription and translation in prokaryotic or eukaryotic organisms. The constructs according to the invention preferably include a promoter 5'-upstream from the respective coding sequence and a terminator sequence 3'-downstream, as well as optionally more of the normal regulatory elements, each one operably linked with the coding sequence. An operable linkage means, for example, the sequential arrangement of promoter, coding sequence, terminator, and optionally more regulatory elements such that each regulatory element can carry out its function in expression of the coding sequence or the antisense sequence as required. This does not necessarily have to be a direct linkage in the chemical sense. Genetic control sequences, such as for example enhancer sequences, can also exercise their function on the target sequence from other DNA molecules.

Examples are sequences to which inducers or repressors bind and thus regulate the expression of the nucleic acid. In addition to these new control sequences or instead of these sequences, the natural regulations of these sequences can still be present in front of the actual structural genes and optionally were genetically modified, so that the natural regulation was switched off and the expression of the gene was increased. But the nucleic acid construct can also be constructed more simply, that is, with no additional regulation signals inserted in front of the above-mentioned genes, and the natural

promoter is not removed with its regulation. Instead, the natural control sequence is mutated so that regulation no longer occurs and gene expression is increased. These modified promoters can also be placed by themselves in front of the natural genes to increase the activity.

The nucleic acid construct can also advantageously contain one or more "enhancer sequences" operably linked with the promoter, which enables increased expression of the nucleic acid sequence. At the 3' end of the DNA sequence, additional sequences can also be advantageously inserted, such as more regulatory elements or terminators. The above-mentioned genes can be contained in the gene construct in one or more copies.

In addition to operable linkage, preferred but not restrictive sequences are further targeting sequences, different from the sequences coding for the transit peptide, to ensure subcellular localization in apoplasts, vacuoles, plastids, the mitochondrion, the endoplasmic reticulum (ER), the cell nucleus, lipid bodies, or other compartments; as well as translation enhancers such as the 5' leader sequence from the tobacco mosaic virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711), and the like.

Control sequences additionally mean sequences that enable homologous or heterologous recombination or insertion into the genome of a host organism or permit removal from the genome. In homologous recombination, for example, the endogenous gene can be entirely inactivated. It also can be exchanged for a synthetic gene with higher and modified activity. Methods such as cre/lox technology permit tissue-specific removal of the target gene from the genome of the host organism under inducible conditions (Sauer B. Methods. 1998; 14(4): 381-92). Here certain flanking sequences are added to the target gene (lox sequences) that later enable removal by means of cre recombinase.

Various control sequences are suitable, depending on the host organism or starting organism, described in more detail below, that is converted to a genetically modified or transgenic organism by introduction of the nucleic acid construct.

Advantageous control sequences for nucleic acid constructs according to the invention, for the vectors according to the invention, for the methods according to the invention for production of Vitamin E and for the genetically modified organisms described below



are contained, for example, in promoters such as *cos-*, *tac-*, *trp-*, *tet-*, *lpp-*, *lac-*, *lpp-lac-*, *lacIq-*, *T7-*, *T5-*, *T3-*, *gal-*, *trc-*, *ara-*, *SP6-*, *l-PR-* or in the *l-PL* promoter, which advantageously are applicable in Gram-negative bacteria.

Further advantageous control sequences are contained in, for example, the Gram-positive promoters *amy* and *SPO2*, the yeast or fungal promoters *ADC1*, *MFa*, *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH*, or the plant promoters *CaMV/35S* [Franck et al., Cell 21(1980) 285-294], *PRP1* [Ward et al., Plant. Mol. Biol. 22 (1993)], *SSU*, *OCS*, *LEB4*, *USP*, *STLS1*, *B33*, *NOS*; *FBPaseP* (WO 98/18940), or the ubiquitin or phaseolin promoter.

In principle, any promoter that can control expression of genes, in particular foreign genes, in plants is suitable as a preferred promoter for the nucleic acid constructs. In particular, a plant promoter or a promoter originating from a plant virus is preferably used. Particularly preferred is the *CaMV 35S* promoter from the cauliflower mosaic virus (Franck et al., Cell 21 (1980), 285-294). This promoter contains various known recognition sequences for transcriptional effectors that all together lead to permanent and constitutive expression of the introduced gene (Benfey et al., EMBO J. 8 (1989), 2195-2202). A further example of a suitable promoter is the legumin B promoter (Accession Number X03677).

Nucleic acid constructs can also contain a chemically inducible promoter, via which the expression of the exogenous gene in the plant can be controlled at a certain point in time. Such promoters, such as for example the *PRP1* promoter (Ward et al., Plant. Mol. Biol. 22 (1993), 361-366), a promoter inducible by salicylic acid (WO 95/19443), a benzenesulfonamide-inducible promoter (EP-A-0388186), a tetracycline-inducible promoter (Gatz et al., (1992) Plant J. 2, 397404), a promoter inducible by abscisic acid (EP-A 335528), or an ethanol- or cyclohexanone-inducible promoter (WO 93/21334) can likewise be used.

In addition, particularly preferred promoters are those that ensure expression in tissues or parts of plants where biosynthesis of Vitamin E or its precursors occurs or where the products advantageously accumulate. Included in particular are promoters for the entire plant based on constitutive expression, such as for example the *CaMV* promoter, the *OCS* promoter from *Agrobacterium* (octopine synthase), the *NOS* promoter from *Agrobacterium* (nopaline synthase), the ubiquitin

promoter, promoters of vacuolar ATPase subunits or the promoter of a proline-rich protein from wheat (WO 91/13991). Included in particular are promoters that ensure leaf-specific expression. Included are the promoter of cytosol FBPase from potato (WO 97/05900), the SSU (small subunit) promoter of rubisco (ribulose-1,5-bisphosphate carboxylase) or the ST-LSI promoter from potato (Stockhaus et al., EMBO J. 8 (1989), 2445-245). Examples of seed-specific promoters are the phaseolin promoter (US 5504200), the USP promoter (Baumlein, H. et al., Mol. Gen. Genet. (1991) 225 (3), 459-467) or the LEB4 promoter (Fiedler, U. et al., Biotechnology (NY) (1995), 13 (10) 1090) together with the LEB4 signal peptide.

Further suitable promoters are, for example, specific promoters for tubers, storage roots or roots, such as for example the Class 1 patatin promoter (B33), the promoter of cathepsin D inhibitor from potato, the promoter of starch synthase (GBSS1) or the sporamin promoter, fruit-specific promoters such as for example the fruit-specific promoter from tomato (EP-A 409625), fruit-ripening-specific promoters such as for example the fruit-ripening-specific promoter from tomato (WO 94/21794), flower-specific promoters such as for example the phytoene synthase promoter (WO 92/16635) or the promoter of the P-rr gene (WO 98/22593) or specific plastid or chromoplast promoters, such as for example the RNA polymerase promoter (WO 97/06250) or also the promoter of phosphoribosyl pyrophosphate amidotransferase from Glycine max (also see GenBank Accession Number U87999) or another node-specific promoter such as in EP-A 249676 [and] can be advantageously used.

In principle, all natural promoters with their regulation sequences as indicated above can be used for the method according to the invention. Moreover, synthetic promoters can also be advantageously used.

Polyadenylation signals that are suitable as control sequences are plant polyadenylation signals, preferably those that essentially correspond to the T-DNA polyadenylation signals from *Agrobacterium tumefaciens*, in particular gene 3 of the T-DNA (octopine synthase) of the Ti plasmid pTiACHS (Gielen et al., EMBO J. 3 (1984), 835 ff) or functional equivalents thereof. Examples of particularly suitable terminator sequences are the OCS (octopine synthase) terminator and the NOS (nopaline synthase) terminator.

A nucleic acid construct is produced, for example, by fusion of a suitable promoter with a suitable anti-HGD, anti-MAAI/FAAH, pro-HG, provitamin E, HGD, MAAI, or FAAH nucleotide sequence, optionally a sequence coding for a transit peptide, preferably a chloroplast-specific transit peptide which preferably is located between the promoter and the respective nucleotide sequence, as well as a terminator or polyadenylation signal. Conventional recombination and cloning techniques are used for this purpose such as are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) as well as in T.J. Silhavy, M.L. Berman and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley Interscience (1987).

As already mentioned, nucleic acid constructs can also be used for which the DNA sequence codes for a pro-HG, provitamin E, HGD, MAAI, or FAAH fusion protein, where one portion of the fusion protein is a transit peptide that controls translocation of polypeptides. An example can be: chloroplast-specific transit peptides which are enzymatically cleaved after translocation into the chloroplasts.

The pro-HG, provitamin E, HGD, MAAI, or FAAH nucleotide sequences are preferably operably linked with the coding sequence of a plant organelle-specific transit peptide. The transit peptide therefore preferably has specificity for individual cell compartments of the plant, for example the plastids, such as for example the chloroplasts, chromoplasts, and/or leucoplasts. The transit peptide guides the expressed polypeptides to the desired target site in the plant, and after it has reached that site it is preferably proteolytically cleaved. The transit peptide coding sequence in the expression construct according to the invention is preferably located 5'-downstream from the pro-HG, provitamin E, HGD, MAAI, or FAAH coding sequence. Included in particular is the transit peptide that is derived from *Nicotiana tabacum* plastid transketolase (TK) or a functional equivalent of this transit peptide (for example, the transit peptide of the small subunit of RubisCO or ferredoxin:NADP oxidoreductase as well as isopentenyl pyrophosphate 2-isomerase).

Another subject matter of the invention relates to transgenic organisms, transformed with at least one nucleic acid construct according to the invention or one vector according to the invention, as well as

cells, cell cultures, tissues, parts (such as, for example, leaves, roots, etc. in plant organisms) or propagation material from such organisms.

Organism, starting organism, or host organism mean prokaryotic or eukaryotic organisms, such as for example microorganisms or plant organisms. Preferred microorganisms are bacteria, yeasts, algae, or fungi.

Preferred bacteria are bacteria of the genera *Escherichia*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes* or cyanobacteria, for example of the genus *Synechocystis*.

Microorganisms are especially preferred which are capable of infecting plants and thereby transferring the constructs according to the invention. Preferred microorganisms are those from the genus *Agrobacterium* and in particular the species *Agrobacterium tumefaciens*.

Preferred yeasts are *Candida*, *Saccharomyces*, *Hansenula* or *Pichia*.

Plant organisms according to the invention are monocotyledonous and dicotyledonous plants. The transgenic plants according to the invention are in particular selected from among monocotyledonous cultivated plants, such as for example cereals such as wheat, barley, millet, rye, triticale, corn, rice, or oats, as well as sugar cane. The transgenic plants according to the invention are also selected in particular from among dicotyledonous cultivated plants, such as for example

Brassicaceae such as rape, cress, arabidopsis, or canola,  
Leguminosae such as soy, alfalfa, pea, bean, or peanut  
Solanaceae such as potato, tobacco, tomato, eggplant, or paprika,  
Asteraceae such as sunflower, tagetes, lettuce, or calendula,  
Cucurbitaceae such as melons, pumpkin, or zucchini,  
as well as linseed, cotton, hemp, flax, red pepper, carrot, sugar beet, and various tree, nut, and vine species.

Particularly preferred are *Arabidopsis thaliana*, *Nicotiana tabacum*, *Tagetes erecta*, *Calendula vulgaris* as well as all genera and species that are suitable for production of oils, such as oilseeds (such as for example rape), nut species, soy, sunflower, pumpkin, and peanut.

Plant organisms according to the invention are also in addition organisms that are photosynthetically active or capable of Vitamin E synthesis, such as for example algae or cyanobacteria, as well as mosses.

Preferred algae are green algae, such as for example algae of the genera *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* or *Dunaliella*.

Transfer of foreign genes into the genome of an organism, for example a plant, is called transformation. The methods described are therefore used for transformation and regeneration of plants from plant tissue or plant cells, for transient or stable transformation. Suitable methods are protoplast transformation by means of polyethylene glycol-induced DNA uptake, the biolistic method with the gene gun, the particle bombardment method, electroporation, incubation of dry embryos in DNA-containing solution, microinjection, and *Agrobacterium*-mediated gene transfer. The indicated methods are described, for example, in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993), 128-143 as well as in Potrykus, Ann. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225). Preferably the construct to be expressed is cloned into a vector that is suitable for transforming *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711).

The efficiency of expression of the transgenetically expressed nucleic acids can be determined, for example, *in vitro* by shoot meristem propagation. Moreover, expression of pro-HG or provitamin E genes that has been modified with respect to type and level, and the effect thereof on the efficiency of Vitamin E biosynthesis, can be tested on test plants in greenhouse experiments.

A further subject matter of the invention is transgenic organisms as described above that are capable of improved Vitamin E production compared with the untransformed wild type.

The invention also relates to cells, cell cultures, parts (such as for example roots, leaves, etc. for transgenic plant organisms), transgenic propagation material, seeds or fruits derived from the transgenic organisms described above.

Improved Vitamin E production in the context of this invention means, for example, the artificially acquired capability for increased efficiency of biosynthesis of at least one compound in the group of tocopherols and tocotrienols, compared with the starting organism that has not been genetically engineered, for the duration of at least one plant generation. In this case, the Vitamin E production in the transgenic organisms, compared with the non-genetically engineered starting organism, is preferably increased by 10%, particularly preferably by 50%, more particularly preferably by 100%. Improved can also mean an advantageously modified qualitative composition of the Vitamin E mixture.

The site of Vitamin E biosynthesis is generally the leaf tissue but also the seeds, so a leaf-specific or seed-specific expression particularly of pro-HG and provitamin E sequences and optionally anti-HGD or anti-MAAI/FAAH sequences is meaningful. However, it is obvious that Vitamin E biosynthesis must not be limited to the seeds, but rather can also occur in all the rests of the parts of the plant in a tissue-specific manner. Moreover, a constitutive expression of the exogenous gene is advantageous. On the other hand, however, an inducible expression is also desirable.

A further subject matter of the invention relates finally to a method for production of Vitamin E that is distinguished by the fact that the desired Vitamin E is isolated in a known way from a culture of a plant organism transformed according to the invention.

Genetically modified plants according to the invention with elevated Vitamin E content, and which can be consumed by humans and animals, can also be used as food or animal feed, for example, directly or after preparation in a known way.

The invention also relates to the use of polypeptides coding for an HGD, MAAI, or FAAH, the genes and cDNAs on which they are based, or nucleic acid constructs according to the invention derived therefrom, vectors according to the invention or organisms according to the invention for production of antibodies, protein- or DNA-binding factors.

The biosynthesis pathway of the HGD-MAAI-FAAH-degradation pathway offers target enzymes for development of inhibitors. Hence the invention also relates to the use of polypeptides coding for an HGD, MAAI, or FAAH, the genes and cDNAs on which they are based, or the nucleic acid constructs according to the invention derived therefrom,

vectors according to the invention or organisms according to the invention as the target for finding inhibitors of HGD, MAAI, or FAAH.

In order to be able to find an efficient inhibitor of HGD, MAAI, or FAAH, it is necessary to have available a suitable test system with which inhibitor — enzyme binding studies can be conducted. In this regard, for example, the complete cDNA sequence of HGD, MAAI, or FAAH is cloned into an expression vector (for example pQE, Qiagen) and overexpressed in *E. coli*. The HGD, MAAI, or FAAH proteins are especially suitable for finding inhibitors that are specific for HGD, MAAI, or FAAH.

Accordingly, the invention relates to a method for finding inhibitors of HGD, MAAI, or FAAH when using the above-mentioned polypeptides, nucleic acids, vectors, or transgenic organisms, wherein the enzyme activity of HGD, MAAI, or FAAH is measured in the presence of a chemical compound and if the enzyme activity decreases compared to the uninhibited activity, then the chemical compound constitutes an inhibitor. For this purpose, HGD, MAAI, or FAAH can be used, for example, in an enzyme test in which the activity of HGD, MAAI, or FAAH is determined in the presence and absence of the active substance under test. From comparison of the two activity determinations, a qualitative and quantitative idea can be obtained about the inhibition behavior of the active substance under test. Using the test system according to the invention, a large number of chemical compounds can be quickly and simply checked for herbicidal properties. The method makes it possible to reproducibly select from a large number of substances specifically those with a stronger effect, in order to subsequently conduct more in-depth tests, familiar to one skilled in the art, with those substances.

Inhibitors of HGD, MAAI, or FAAH are suitable for increasing Vitamin E biosynthesis in a way functionally analogous to what is described above for anti-HGD or anti-MAAI/FAAH nucleic acid sequences. Hence a further subject matter of the invention is to improve Vitamin E production by using inhibitors of HGD, MAAI, or FAAH. Improved production of Vitamin E can have a positive effect on the plant, since these compounds have an important function in protecting the plant from harmful environmental influences (solar radiation, oxygen radicals). In this respect, an increase in Vitamin E production can act as a growth promoter. Hence a further subject matter of the invention is use of inhibitors of HGD, MAAI, or

FAAH, obtained by the method described above, as growth regulators.

#### Sequences

- SEQ ID NO. 1: Homogentisate 1,2-dioxygenase (HGD) gene from *Arabidopsis thaliana*
- SEQ ID NO. 2: Homogentisate 1,2-dioxygenase (HGD) cDNA from *Arabidopsis thaliana*
- SEQ ID NO. 3: Homogentisate 1,2-dioxygenase (HGD) polypeptide from *Arabidopsis thaliana*
- SEQ ID NO. 4: Fumarylacetoacetate hydrolase (FAAH) cDNA from *Arabidopsis thaliana*
- SEQ ID NO. 5: Fumarylacetoacetate hydrolase (FAAH) polypeptide from *Arabidopsis thaliana*
- SEQ ID NO. 6: Maleylacetoacetate isomerase (MAAI) gene from *Arabidopsis thaliana*
- SEQ ID NO. 7: TyrA gene coding for a bifunctional chorismate mutase/ prephenate dehydrogenase
- SEQ ID NO. 8: TyrA polypeptide coding for a bifunctional chorismate mutase/ prephenate dehydrogenase
- SEQ ID NO. 9: Fumarylacetoacetate hydrolase (FAAH) gene from *Arabidopsis thaliana*,
- SEQ ID NO. 10: Hydroxyphenyl pyruvate dioxygenase (HPPD) cDNA from *Arabidopsis thaliana*
- SEQ ID NO. 11: Hydroxyphenyl pyruvate dioxygenase (HPPD) polypeptide from *Arabidopsis thaliana*.
- SEQ ID NO. 12: Homogentisate 1,2-dioxygenase (HGD) cDNA fragment from *Brassica napus*
- SEQ ID NO. 13: Homogentisate phytyltransferase cDNA from *Synechocystis* PCC6803
- SEQ ID NO. 14: Homogentisate phytyltransferase polypeptide from *Synechocystis* PCC6803
- SEQ ID NO. 15: Artificial cDNA optimized for codon usage, coding for hydroxyphenyl pyruvate dioxygenase (HPPDop) from *Streptomyces avermitilis*
- SEQ ID NO. 16: Hydroxyphenyl pyruvate dioxygenase polypeptide, from *Streptomyces avermitilis*
- SEQ ID NO. 17: Maleylacetoacetate isomerase (MAAI) cDNA from *Arabidopsis thaliana*
- SEQ ID NO. 18: Maleylacetoacetate isomerase (MAAI) polypeptide from *Arabidopsis thaliana*
- SEQ ID NO. 19:  $\gamma$ -Tocopherol methyltransferase cDNA from *Arabidopsis thaliana*
- SEQ ID NO. 20:  $\gamma$ -Tocopherol methyltransferase polypeptide from *Arabidopsis thaliana*



- SEQ ID NO. 21: 3-Methyl-6-phytylhydroquinone methyltransferase cDNA from *Synechocystis* PCC6803
- SEQ ID NO. 22: 3-Methyl-6-phytylhydroquinone methyltransferase polypeptide from *Synechocystis* PCC6803
- SEQ ID NO. 23: Geranylgeranyl pyrophosphate oxidoreductase cDNA from *Nicotiana tabacum*.
- SEQ ID NO. 24: Geranylgeranyl pyrophosphate oxidoreductase polypeptide from *Nicotiana tabacum*.
- SEQ ID NO. 25: Primer (5'-HGD *Brassica napus*)  
5'-GTCGACGGNCCNATNGGNGCNAANGG-3'
- SEQ ID NO. 26: Primer (3'-NOS Terminator)  
5'-AAGCTTCCGATCTAGTAACATAGA-3'
- SEQ ID NO. 27: Primer (5'-35 S Promoter)  
5'-ATTCTAGACATGGAGTCAAAGATTCAAATAGA-3'
- SEQ ID NO. 28: Primer (3'-OCS terminator)  
5'-ATTCTAGAGGACAATCAGTAAATTGAACGGAG-3'
- SEQ ID NO. 29: Primer (5'-MAAI *A. thaliana*)  
5'-atgtcgacATGTCTTATGTTACCGAT-3'
- SEQ ID NO. 30: Primer (3'-MAAI *A. thaliana*)  
5'-atggatccCTGGTTCATATGATACA-3'
- SEQ ID NO. 31: Primer (5'-FAAH *A. thaliana*)  
5'-atgtcgacGAAACTCTGAACCATAT-3'
- SEQ ID NO. 32: Primer (3'-FAAH *A. thaliana*)  
5'-atggtaccGAATGTGATGCCTAAGT-3'
- SEQ ID NO. 33: Primer (3'-HGD *Brassica napus*)  
5'-GGTACCTCRAACATRAANGCCATNGTNCC-3'
- SEQ ID NO. 34: Primer (5'-legumin promoter)  
5'-GAATTCGATCTGTCGTCTCAAACCTC-3'
- SEQ ID NO. 35: Primer (3'-legumin promoter)  
5'-GGTACCGTGATAGTAAACAACCTAATG-3'
- SEQ ID NO. 36: Primer (5'-transit peptide)  
5'-ATGGTACCTTTTTTGCATAAACTTATCTTCATAG-3'
- SEQ ID NO. 37: Primer (3'-transit peptide)  
5'-  
ATGTCGACCCGGGATCCAGGGCCCTGATGGGTCCCATTTC  
C-3'
- SEQ ID NO. 38: Primer (5'-NOS terminator)  
5'-GTCGACGAATTTCCCCGAATCGTTC-3'
- SEQ ID NO. 39: Primer (3'-NOS terminator II)  
5'-AAGCTTCCGATCTAGTAACATAGA-3'
- SEQ ID NO. 40: Primer (5'-legumin promoter II)  
5'-AAGCTTGATCTGTCGTCTCAAACCTC-3'
- SEQ ID NO. 41: Maleylacetoacetate isomerase (MAAI) gene (fragment) from *Arabidopsis thaliana*
- SEQ ID NO. 42: Fumarylacetoacetate hydrolase (FAAH) gene (fragment) from *Arabidopsis thaliana*
- SEQ ID NO. 43: Primer (5'-35 S promoter)  
5'-ATGAATTCCATGGAGTCAAAGATTCAAATAGA-3'

SEQ ID NO. 44: Primer (3'-OCS terminator)  
5'-ATGAATTCGGACAATCAGTAAATTGAACGGAG-3'

### Examples

The invention is explained in more detail in the following exemplified embodiments with reference to the attached figures. Abbreviations with the following meaning are used for this purpose:

A = 35S-Promoter	B = HGD in antisense orientation
C = OCS Terminator	D = Legumin B promoter
E = Transit peptide of FNR	F = HPPDop (HPPD with optimized codon usage)
G = NOS-Terminator	H = MAA1 in antisense orientation
I = FAAH in antisense orientation	

The direction of the arrows in the figures indicates in each case the reading direction for the corresponding genes.

### List of figures:

- Figure 1 a flowsheet for the Vitamin E biosynthesis pathway in plants;
- Figure 2 Schematic representation of the anti-HGD coding plasmids pBinARHGDanti (I) and pCRScriptHGDanti (II) ;
- Figure 3 Schematic representation of the HPPDop coding plasmids pUC19HPPDop (III) and pCRScriptHPPDop (IV);
- Figure 4 Schematic representation of the transformation vectors pPTVHGDanti (V) and the bifunctional transformation vector pPTV HPPDop HGD anti (VI), which expresses HPPDop in seeds of transformed plants and simultaneously suppresses expression of endogenous HGD.
- Figure 5 Schematic representation of the transformation vector pPZP200HPPDop (VII).
- Figure 6 Schematic representation of the transformation vectors pGEMT MAA1 anti (VIII) and pBinAR MAA1 anti (IX)
- Figure 7 Schematic representation of the transformation vectors pCRScript MAA1 anti (X) and pZPNBN MAA1 anti (XI)
- Figure 8 Schematic representation of the transformation vectors [sic] pGEMT FAAH anti (XII)

Figure 9 Schematic representation of the transformation vectors pBinAR FAAH anti (XIII) and pZPNBN FAAH anti (XIV)

General Methods:

Chemical synthesis of oligonucleotides can, for example, be carried out in a known way by the phosphoramidite method (Voet and Voet, 2nd Edition, Wiley Press New York, pp. 896-897). The cloning steps conducted in the context of this invention, such as for example restriction digests, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linking DNA fragments, transformation of *E. coli* cells, cultivation of bacteria, phage propagation, and sequence analysis of recombinant DNA were carried out as described in Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6. Sequencing of the recombinant DNA molecules was performed with a laser fluorescence DNA sequencer from Licor (sales through MWG Biotech, Ebersbach) according to Sanger's method (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Example 1

Cloning hydroxyphenylpyruvate dioxygenase (HPPD) with DNA sequence optimized for expression in *Brassica napus*

The amino acid sequence of hydroxyphenylpyruvate dioxygenase (HPPD) from *Streptomyces avermitilis* (Accession Number U11864, SEQ ID NO:16) was back-translated to a DNA sequence, taking into consideration the codon usage in *Brassica napus* (rape). The codon usage was determined from the databank at <http://www.dna.affrc.go.jp/~nakamura/index.html>. The derived sequence was synthesized (SEQ ID NO:15), with attachment of Sall cleavage sites, by ligation of overlapping oligonucleotides followed by PCR amplification (Rouwendaal, GJA et al. (1997) PMB 33: 989-999). The correctness of the sequence of the synthetic gene was checked by sequencing. The synthetic gene was cloned into the vector pBluescript II SK+ (Stratagene). (This codon-optimized sequence is also called HPPDop in the following.)

## Example 2:

Cloning a homogentisate dioxygenase (HGD) from *Brassica napus*

a ) Isolation of total RNA from flowers of *Brassica napus*

Open flowers of *Brassica napus* var. Westa [*sic, should be Westar*] were harvested and frozen in liquid nitrogen. The material was then pulverized in a mortar and taken up in Z6-Puffer (8 M guanidinium hydrochloride, 20 mM MES, 20 mM EDTA, adjusted to pH 7.0 with NaOH; mixed with 400 ml mercaptoethanol/100 ml Puffer immediately before use). The suspension was then transferred to a reaction vessel and shaken with one volume of phenol/chloroform/isoamyl alcohol 25:24:1. After 10 minutes of centrifuging at 15000 rpm, the supernatant was transferred to a new reaction vessel and the RNA was precipitated with 1/20 volume of 1 N acetic acid and 0.7 volumes of ethanol (absolute). After centrifuging again, the pellet was washed first in 3 M sodium acetate solution and then, after another centrifuging, in 70% ethanol. Then the pellet was dissolved in DEPC-treated water (diethylpyrocarbonate) and the RNA concentration was determined photometrically.

b) Preparation of cDNA from total RNA from flowers of *Brassica napus*

20 mg of total RNA were first mixed with 3.3 ml of 3 M sodium acetate solution, 2 ml of 1 M magnesium sulfate solution and brought up to a final volume of 10 ml with DEPC-treated water. Then 1 ml RNase-free DNase (Boehringer Mannheim) was added and it was incubated for 45 min at 37 degrees. After removal of the enzymes by shaking with phenol/chloroform/isoamyl alcohol, the RNA was precipitated with ethanol and the pellet was taken up in 100 ml DEPC-treated water. 2.5 mg RNA from this solution was transcribed into cDNA using a cDNA-Kit (Gibco BRL) according to the manufacturer's instructions.

c) PCR-amplification of a partial fragment of HGD from *Brassica napus*

By comparing the DNA sequences for known homogentisate dioxygenases (HGD) from *Arabidopsis thaliana* (Accession Number U80668), *Homo sapiens* (Accession Number U63008) and *Mus musculus* (Accession Number U58988), an oligonucleotide was derived for a PCR with an Sall restriction cleavage site added to the 5' end and an Asp718 site added to the 3' end. The oligonucleotide at the 5'-end comprises the sequence:

5'-GTCGACGGNCCNATNGGNGCNAANGG-3' (SEQ ID NO:25),

beginning with base 661 of the *Arabidopsis* gene. The oligonucleotide at the 3'-end comprises the sequence:

5'-GGTACCTCRAACATRAANGCCATNGTNCC-3' (SEQ ID NO:33),

beginning with base 1223 of the *Arabidopsis* gene, where N means inosine and R stands for insertion of A or G into the oligonucleotide.

The PCR reaction was carried out with *Taq* polymerase from TAKARA according to the manufacturer's instructions. 0.3 mg cDNA was used as a template. The PCR program was:

- 1 cycle with: 94°C (1 min)  
5 cycles with: 94°C (4 sec), 50°C (30 sec), 72°C (1 min)  
5 cycles with: 94°C (4 sec), 48°C (30 sec), 72°C (1 min)  
25 cycles with: 94°C (4 sec), 46 degrees (30 sec), 72 degrees (1 min)  
1 cycle with: 72 degrees (30 min)

The fragment was purified using NucleoSpin Extract (Machery and Nagel) and cloned into the vector pGEMT (Promega) according to the manufacturer's instructions. The correctness of the fragment was checked by sequencing.

Example 3: Preparation of a plant transformation construct for overexpression of HPPD with optimized DNA sequence (HPPDop) and switching off HGD

To produce plants that express HPPDop in seeds and in which the expression of endogenous HGD is suppressed by an antisense technique, a binary vector was prepared that contains both gene sequences (Figure 4, Construct VI).

#### a) Preparation of an HPPDop nucleic acid construct

For this purpose, first the components of the HPPDop expression cassette, consisting of legumin B promoter (Accession Number X03677), the transit peptide of ferredoxin:NADP+ oxidoreductase from spinach (FNR; Jansen, T. et al (1988) Current Genetics 13, 517-522) and the NOS terminator (contained in pBI101, Accession Number U12668) were provided with the required restriction cleavage sites by PCR.

The legumin promoter was amplified by PCR from the plasmid plePOCS (Bäumlein, H. et al. (1986) Plant J. 24, 233-239) with the upstream oligonucleotide:

5'-GAATTCGATCTGTCGTCTCAAACCTC-3' (SEQ ID NO: 34)

and the downstream oligonucleotide:

5'-GGTACCGTGATAGTAAACAATAATG-3' (SEQ ID NO: 35)

and cloned into the vector PCR-Script (Stratagene) according to the manufacturer's instructions.

The transit peptide was amplified by PCR from plasmid pSK-FNR (Andrea Babette Regierer, Molecular Genetics Approaches to Modifying Phosphate Utilization Efficiency of Higher Plants, P+H Wissenschaftlicher Verlag, Berlin 1998 ISBN: 3-9805474-9-3) with the 5' oligonucleotide:

5'-ATGGTACCTTTTTTGCATAAACTTATCTTCATAG-3' (SEQ ID NO: 36)

and the 3' oligonucleotide:

5'-ATGTCGACCCGGGATCCAGGGCCCTGATGGGTCCCATTTCCTCC-3'  
(SEQ ID NO: 37)

The NOS terminator was amplified by PCR from the plasmid pBI101 (Jefferson, R.A. et al (1987) EMBO J. 6 (13), 3901-3907) with the 5'-oligonucleotide:

5'-GTCGACGAATTTCCCCGAATCGTTC-3' (SEQ ID NO: 38)

and the 3' oligonucleotide:

5'-AAGCTTCCGATCTAGTAACATAGA-3' (SEQ ID NO: 26)

The amplicon was cloned into the vector pCR-Script (Stratagene) according to the manufacturer's instructions.

For the nucleic acid construct, first the NOS terminator was subcloned as an Sall/HindIII fragment into an appropriately cut pUC19 vector (Yanisch-Perron, C., et al (1985) Gene 33, 103-119). Then the transit peptide was inserted into this plasmid as an Asp718/Sall fragment. The legumin promoter

was then cloned in as an *EccRI*/*Asp718* fragment. The gene *HPPDop* was inserted in this construct as an *Sall* fragment (Figure 3, Construct III).

The finished cassette in *pUC19* was used as a template for a PCR, where the oligonucleotide

5'-AAGCTTGATCTGTCGTCTCAAACCTC-3' (SEQ ID NO: 40)

was used for the legumin promoter, and the oligonucleotide

5'-AAGCTTCCGATCTAGTAACATAGA-3' (SEQ ID NO: 39)

was used for the NOS terminator. The amplicon was cloned into *pCR-Script* and named *pCR-ScriptHPPDop* (Figure 3, Construct IV).

d) Preparation of an antiHGD nucleic acid construct

To switch off HGD with an antisense technique, the gene fragment was cloned as an *Sall*/*Asp718* fragment into the vector *pBinAR* (Höfgen, R. and Willmitzer, L. (1990) *Plant Sci.* 66: 221-230), which contained the 35S promoter and the OCS terminator (Figure 2, Construct I). The construct served as a template for a PCR reaction with the oligonucleotide:

5'-ATTCTAGACATGGAGTCAAAGATTCAAATAGA-31 (SEQ ID NO: 27),

specifically for the 35S promoter sequence;

and the oligonucleotide:

5'-ATTCTAGAGGACAATCAGTAAATTGAACGGAG-3' (SEQ ID NO: 28).

specifically for the OCS terminator sequence.

The amplicon was cloned into the vector *pCR-Script* (Stratagene) and named *pCRScriptHGDanti* (Figure 2, Construct II).

c) Preparation of binary vectors

To construct a binary vector for rape transformation, first the construct *HGDanti* from *pCRScriptHGDanti* was cloned as an *XbaI* fragment into the vector *pPTV* (Becker, D., (1992) *PMB* 20, 1195-1197) (Figure 4, Construct V). The construct *LegHPPDop* from *pCRScriptHPPDop* was inserted into this plasmid as an *HindIII*

fragment. This plasmid was named pPTVHPPDopHGDanti (Figure 4, Construct VI),

**Example 4:**

Preparation of constructs for cotransformation for overexpression of HPPDop and switching off HGD in *Brassica napus* plants

For cotransformation of plants with HPPDop and antiHGD, the construct legumin B promoter/transit peptide/HPPDop/NOS from the vector pCRScriptHPPDop (Figure 3, Construct IV) was excised as an HindIII fragment and inserted into the appropriately cut vector pPZP200 (Hajdukiewicz, P. et al., (1994) PMB 25(6): 989-94) (Figure 5, Construct VII). Later this plasmid was used for cotransformation of plants together with the vector pPTVHGDanti (Figure 4, Construct V) from Example 3c.

**Example 5:**

Cloning a genomic fragment of maleylacetoacetate isomerase from *Arabidopsis thaliana*

a) Isolation of genomic DNA from leaves of *A. thaliana*:

The extraction buffer used had the following composition:

- 1 Volume DNA extraction buffer (0.35 M sorbitol, 0.1 M Tris, 5 mM EDTA, pH 8.25 HCl)
- 1 Volume Nuclei Lysis buffer (0.2 M Tris-HCl, pH 8.0, 50 mM EDTA, 2 M NaCl, 2% Hexadecyltrimethylammonium bromide (CTAB))
- 0.4 Volumes 5% sodium sarkosyl
- 0.38 g/100 ml sodium bisulfite

100 mg of leaf material of *A. thaliana* was harvested and frozen in liquid nitrogen. The material was then pulverized in a mortar and taken up in 750 µl extraction buffer. The mixture was heated for 20 min at 65°C and then shaken with one volume of chloroform/isoamyl alcohol (24:1). After centrifuging for 10 min at 10000 rpm in an Heraeus Picofuge, the supernatant was mixed with one volume of isopropanol and the DNA precipitating was pelletized again for 5 minutes at 10000 rpm. The pellet was washed in 70% ethanol, dried at room temperature for 10 min, and then dissolved in 100 µl TE-RNase



buffer (10 mM Tris HCl pH 8.0, 1 mM EDTA pH 8.0, 100 mg/l RNase).

b) Cloning the gene for MAAI from *Aribidopsis thaliana*

Using the protein sequence of mouse MAAI (*Mus musculus*), a BLAST search in the NCBI databank (<http://www.ncbi.nlm.nih.gov/BLAST/>) was used to identify the MAAI gene from *A. thaliana* (GenBank Accession Number AAC78520.1). The sequence is annotated in GenBank as a putative glutathione S-transferase. Using the ID numbers of the protein sequence, the corresponding DNA sequence could be determined and the oligonucleotides derived. The oligonucleotides were added respectively to the 5' end, an Sall restriction cleavage site, and to the 3' end, a BamHI cleavage site. The oligonucleotide at the 5'-end comprises the sequence:

5'-atgtcgacATGTCTTATGTTACCGAT-3' (SEQ ID NO: 29)

beginning with base 37 of the cDNA, the first codon, the oligonucleotide at the 3' end comprises the sequence

5'-atggatccCTGGTTCATATGATACA-3' (SEQ ID NO: 30)

beginning with base pair 803 of the cDNA sequence. The PCR reaction was carried out with *Taq* polymerase (Manufacturer: TaKaRa Shuzo Co. Ltd.) according to the manufacturer's instructions. The mixture had the following composition: 10 µl buffer (20 mM Tris-HCl pH 8.0, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5% Tween 20, 0.5 % Nonidet P-40, 50% glycerol), 100 pmol each of the two oligonucleotides, 20 nM each of dATP, dCTP, dGTP, dTTP, 2.5 units *Taq* polymerase, 1 µg genomic DNA, distilled water to 100 µl. The PCR program was:

- 5 cycles with: 94°C (4 sec), 52°C (30 sec), 72°C (1 min)
- 5 cycles with: 94°C (4 sec), 50°C (30 sec), 72°C (1 min)
- 25 cycles with: 94°C (4 sec), 48°C (30 sec), 72°C (1 min)

The amplified fragment (SEQ ID NO: 41) was purified using NucleoSpin Extract (Machery-Nagel) and was cloned into the vector pGEMTeasy from Promega according to the manufacturer's instructions (Figure 6, Construct VIII). The correctness of the fragment was checked by sequencing. Using the restriction cleavage sites added to the sequence by the primer, the gene was cloned as an Sall/BamHI fragment into the appropriately cut vector BinAR (Höfgen, R. and Willmitzer, L., (1990) Plant Sci. 66: 221-230) (Figure 6, Construct IX). This contains the 35S promoter of the cauliflower mosaic virus

and the OCS termination sequence. The construct served as a template for a PCR reaction with the oligonucleotide

5'-ATGAATTCCATGGAGTCAAAGATTCAAATAGA-3' (SEQ ID NO: 43)

specifically for the 35S promoter sequence, and with the oligonucleotide

5'-ATGAATTCGGACAATCAGTAAATTGAACGGAG-3' (SEQ ID NO: 44)

specifically for the OCS terminator. An EcoRI recognition sequence was added to both oligonucleotides. The PCR was carried out with *Pfu* Polymerase (Manufacturer: Stratagene). The mixture had the following composition: 10 µl buffer (200 mM Tris HCl, pH 8.8, 20 mM MgSO<sub>4</sub>, 100 mM KCl, 100 mM Ammonium sulfate, 1% Triton X-100, 1 g/l nuclease-free BSA), 100 pmol each of the two oligonucleotides, 20 nM each of dATP, dCTP, dGTP, dTTP, 2.5 units *Pfu* polymerase, 1 ng plasmid DNA, distilled water to 100 µl. The PCR program was:

- 5 cycles with: 94°C (4 sec), 52°C (30 sec), 72°C (2 min)
- 5 cycles with: 94°C (4 sec), 50°C (30 sec), 72°C (2 min)
- 25 cycles with: 94°C (4 sec), 48°C (30 sec), 72°C (2 min)

The PCR fragment was purified using NucleoSpin Extract (Machery-Nagel) and was cloned into the vector pCR-Script (Stratagene) (Figure 7, Construct X).

#### Example 6: Constructing binary vectors

To construct a binary vector for *Arabidopsis* and rape transformation, the construct from the vector PCR-Script was cloned as an EcoRI fragment into the vector pZPNBN. pZPNBN is a pPZP200 derivative (Hajdukiewicz, P. et al., (1994) PMB 25(6): 989-94), to which phosphinotricin resistance under the control of the NOS promoter had been added previously in front of the NOS terminator. (Figure 7, Construct XI)

#### Example 7. Cloning a genomic fragment of fumarylacetoacetate isomerase from *Arabidopsis thaliana*

A BLAST search was carried out using the protein sequence of FAAH from *Emericella nidulans*, and a protein sequence from *A. thaliana* was identified with 59% homology. FAAH from *A. thaliana* has Accession Number AC002131. Using the ID Number of the protein sequence, the DNA sequence could be determined and the oligonucleotides derived.

An Sall restriction cleavage site was added to the 5' oligonucleotide and an Asp718 cleavage site was added to the 3' oligonucleotide. The oligonucleotide at the 5' end of FAAH comprises the sequence:

5'-atgtcgacGGAACTCTGAACCATAT-3' (SEQ ID NO: 31)

beginning with base 40258 of BAC F12F1, the oligonucleotide at the 3' end comprises the sequence

5'-atggtaccGAATGTGATGCCTAAGT-3' (SEQ ID NO: 32)

beginning with base pair 39653 of BAC. The PCR reaction was carried out with *Taq* polymerase (Manufacturer: TaKaRa Shuzo Co. Ltd.). The mixture had the following composition: 10 µl buffer (20 mM Tris-HCl pH 8.0, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5% Tween 20, 0.5 % Nonidet P-40, 50% glycerol), 100 pmol each of the two oligonucleotides, 20 nM each of dATP, dCTP, dGTP, dTTP, 2.5 units *Taq* polymerase, 1 µg genomic DNA, distilled water to 100 µl. The PCR program was:

- 5 cycles with: 94°C (4 sec), 52°C (30 sec), 72°C (1 min)
- 5 cycles with: 94°C (4 sec), 50°C (30 sec), 72°C (1 min)
- 25 cycles with: 94°C (4 sec), 48°C (30 sec), 72°C (1 min)

The fragment ((SEQ ID NO: 42) was purified using NucleoSpin Extract (Machery-Nagel) and was cloned into the vector pGEMTeasy from Promega according to the manufacturer's instructions (Figure 8, Construct XII).

The correctness of the fragment was checked by sequencing. Using the restriction cleavage sites added to the sequence by the primer, the gene was cloned as an Sall/Asp718 fragment into the appropriately cut vector BinAR (Höfgen, R. and Willmitzer, L., Plant Sci. 66: 221-230, 1990). This contains the 35S promoter of the cauliflower mosaic virus and the OCS termination sequence (Fig. 9, Construct XIII).

To construct a binary vector for *Arabidopsis* and rape transformation, the construct from the vector pBinAR was cloned as an EcoRI/HindIII fragment into the vector pZPNBN. pZPNBN is a pPZP200 derivative ((Hajdukewicz, P. et al., Plant Molecular Biology, 25; 989-994, 1994), to which phosphinotricin resistance, under the control of the NOS promoter, had been added previously in front of the NOS terminator.

## Example 8:

Preparation of transgenic *Arabidopsis thaliana* plants

Wild-type *Arabidopsis thaliana* plants (cv. Columbia) were transformed with the *Agrobacterium tumefaciens* strain (EHA105) based on a modified vacuum infiltration method of Clough and Bent (Clough, S. and Bent A., Plant J. 16 (6) ; 735-43, 1998) and of Bechtold et al. (Bechtold, N. et al., C.R. Acad Sci Paris. 1144(2): 204-212, 1993). The *Agrobacterium tumefaciens* cells used had been transformed previously with the plasmids pZPNBN-MAAlanti or pZPNBN-FAAHanti.

Seeds of the primary transformants were screened for phosphinotricin resistance by planting the seeds and spraying the embryos with the herbicide Basta (phosphinotricin). There were isolated Basta-resistant embryos, and as fully developed plants they were used for biochemical analysis.

Example 9: Preparation of transgenic rape (*Brassica napus*) plants

Preparation of transgenic rape plants was guided by a protocol of Bade, J.B. and Damm, B. (Bade, J.B. and Damm, B. (1995) in: Gene Transfer to Plants, Potrykus, I. and Spangenberg, G., eds, Springer Lab Manual, Springer Verlag, 1995, 30-38), where the composition of the media and buffer used is also given.

Transformation was conducted with *Agrobacterium tumefaciens* strain EHA105. Either plasmid pPTVHPPDopHGDanti (Figure 4, Construct VI) or, after cultivation of mixed cultures of *Agrobacteria* with plasmids pPTVHGDanti (Figure 4, Construct V) and pPZP200HPPDop (Figure 5, Construct VII) [those plasmids?] were used. The outer surface of seeds of *Brassica napus* var. Westar were sterilized with 70% Ethanol (v/v), washed for 10 minutes at 55°C in water, in 1% hypochlorite solution (25% v/v Teepol, 0.1 % v/v Tween 20), incubated for 20 minutes, and washed six times with sterile water (for 20 minutes each). The seeds were dried on filter paper for three days, and 10-15 seeds were germinated in a glass flask with 15 ml germination medium. The roots and apices were removed from several embryos (about 10 cm long) and the remaining hypocotyls were cut into pieces about 6 mm long. Approximately 600 explants obtained in this way were washed for 30 minutes with 50 ml basal medium and transferred to a 300 ml flask. After addition of 100 ml callus induction medium, the cultures were incubated for 24 hours at 100 rpm.

Overnight cultures of the *Agrobacterium* strains were set up at 29°C in Luria Broth Medium with kanamycin (20 mg/l), of which 2 ml in 50 ml Luria Broth Medium without kanamycin was incubated for 4 hours at 29°C until a value of OD<sub>600</sub> [optical density or absorbance at 600 nm] = 0.4-0.5 was achieved. After pelletizing the culture at 2000 rpm for 25 min, the cell pellets were suspended again in 25 ml of basal medium. The concentration of bacteria in the solution was adjusted by adding more basal medium until a value of OD<sub>600</sub> = 0.3 was achieved. Equal parts of the solutions of both strains were mixed for cotransformation.

The callus induction medium was removed from the rape explants with sterile pipets, 50 ml *Agrobacterium* solution was added, then it was carefully mixed and incubated for 20 min. The *Agrobacteria* suspension was removed, the rape explants were washed for 1 min with 50 ml of callus induction medium and then 100 ml of callus induction medium was added. Co-cultivation was conducted for 24 h on a rotary shaker at 100 rpm. Co-cultivation was stopped by removing the callus induction medium, and the explants were washed twice for 1 minute each with 25 ml wash medium and twice for 60 min with 100 ml wash medium each at 100 rpm. The wash medium with the explants was transferred to 15 cm petri dishes and the medium was removed with sterile pipets.

For the regeneration, 20-30 explants each were transferred to 90 mm petri dishes containing 25 ml sprout induction medium with phosphinotricin. The petri dishes were sealed with 2 layers of Leukopor [sealing tape] and incubated at 25°C and 2000 lux with a light cycle of 16 hours light / 8 hours dark. Every 12 days, the developing calluses were moved to fresh petri dishes with sprout induction medium. All further steps for regeneration of the entire plants were carried out as described by Bade, J.B and Damm, B. (in: Gene Transfer to Plants, Potrykus, I. and Spangenberg, G., eds, Springer Lab Manual, Springer Verlag, 1995, 30-38).

#### Example 10: Study of transgenic plants

In order to confirm that Vitamin E biosynthesis in transgenic plants is affected by inhibition of HGD, MAAI, and/or FAAH, the tocopherol and tocotrienol contents in leaves and seeds were analyzed for plants transformed with the described constructs (*Arabidopsis thaliana*, *Brassica napus*). For this purpose, the transgenic plants were cultivated in the greenhouse and plants expressing antisense RNA of HGD, MAAI, and/or FAAH were studied by Northern Blot analysis.

The tocopherol content and the tocotrienol content was determined in the leaves and seeds of these plants. Digestion of the plant materials was carried out by incubating three times in the Eppendorf shaker at 30°C, 1000 rpm in 100% methanol for 15 minutes, where each time the supernatants obtained were purified. Further incubation steps did not result in any further release of tocopherols or tocotrienols. In order to avoid oxidation, the extracts obtained were analyzed immediately after extraction using a Waters Alliance [*sic, should be Alliance*] 2690 HPLC system. Tocopherols and tocotrienols were separated on a reverse phase column (ProntoSil 200-3-C30, Bischoff) with 100% methanol as the mobile phase and were identified based on Standards (Merck). The detection system was the fluorescence of the substances (excitation 295 nm, emission 320 nm) that was determined using a Jasco FP 920 Fluorescence Detector.

In all cases, the tocopherol or tocotrienol concentration was elevated, compared with untransformed plants, in transgenic plants that additionally expressed a nucleic acid according to the invention.

## Claims

1. A method for Vitamin E formation by influencing Vitamin E biosynthesis, wherein homogentisate degradation is reduced by decreasing homogentisate 1,2-dioxygenase (HGD) activity, maleylacetoacetate isomerase (MAAI) activity, and/or fumarylacetoacetate hydrolase (FAAH) activity.
2. The method according to Claim 1, wherein the MAAI activity and/or the FAAH activity is reduced and at the same time
  - a) conversion of homogentisate to Vitamin E is improved, or
  - b) biosynthesis of homogentisate is improved.
3. The method according to Claim 1, wherein the HGD activity is reduced and at the same time
  - a) conversion of homogentisate to Vitamin E is improved, or
  - b) the TyrA gene is overexpressed.
4. A method for increased formation of Vitamin E by influencing Vitamin E biosynthesis, wherein
  - a) conversion of homogentisate to Vitamin E and at the same time
  - b) biosynthesis of homogentisateare improved.
5. The method according to Claims 1 to 3, wherein inhibitors of MAAI, HGD, or FAAH are used in cultivation of a plant organism.

Drawings.  
Sequences.

**REPLACEMENT SHEET (RULE 26)**

6. A nucleic acid construct, containing a nucleic acid sequence (anti-MAAI/FAAH) capable of reducing the MAAI activity or the FAAH activity, or a functional equivalent thereof.
7. A nucleic acid construct according to Claim 6, additionally containing
  - a) a nucleic acid sequence (pro-HG) capable of increasing homogentisate biosynthesis, or a functional equivalent thereof; or
  - b) a nucleic acid sequence (provitamin E) capable of increasing Vitamin E biosynthesis starting from homogentisate, or a functional equivalent thereof; or
  - c) a combination of a) and b).
8. A nucleic acid construct, containing a nucleic acid sequence (anti-HGD) capable of inhibition of HGD, or for a functional equivalent thereof.
9. A nucleic acid construct according to Claim 8, additionally containing
  - a) a nucleic acid sequence coding for bifunctional chorismate mutase — prephenate dehydrogenase enzyme (TyrA) or a functional equivalent thereof; or
  - b) a nucleic acid sequence (provitamin E) capable of increasing Vitamin E biosynthesis starting from homogentisate, or a functional equivalent thereof; or
  - c) a combination of a) and b).
10. A nucleic acid construct containing a nucleic acid sequence (pro-HG) capable of increasing biosynthesis of homogentisate (HG), or a functional equivalent thereof, and at the same time a nucleic acid sequence (provitamin E) capable of increasing Vitamin E biosynthesis starting from homogentisate, or a functional equivalent thereof.
11. A nucleic acid construct according to Claims 6 to 10, containing an anti-MAAI/FAAH sequence or an anti-HGD sequence, that



- a) can be transcribed to an antisense nucleic acid sequence capable of inhibiting MAAI/ FAAH activity or HGD activity, or
  - b) causes inactivation of MAAI/FAAH or HGD by homologous recombination, or
  - c) codes for a binding factor that binds MAAI/FAAH or HGD to the genes and therefore decreases transcription of those genes.
12. A nucleic acid construct according to Claims 7 and 10, containing a pro-HG sequence selected from the genes coding for an HPPD, TyrA.
13. A nucleic acid construct according to Claims 7, 9, and 10, containing a provitamin E sequence selected from the genes coding for an HPGT, geranylgeranyl oxidoreductase, 2-methyl-6-phytylplastoquinol methyltransferase,  $\gamma$ -tocopherol methyltransferase.
14. A recombinant vector, containing
- a) a nucleic acid construct according to any one of Claims 6 to 13; or
  - b) a nucleic acid coding for an HGD, MAAH [*sic, typo for MAAI*], or FAAH, as well as functional equivalents thereof, or
  - c) a combination of options a) and b).
15. A recombinant vector according to Claim 14, wherein the nucleic acids or nucleic acid constructs are operably linked with a genetic control sequence and that is capable of transcription of sense- or antisense-RNA.
16. A transgenic organism, transformed with a nucleic acid construct according to Claims 6 to 13, or a recombinant vector according to Claims 14 or 15.
17. A transgenic organism according to Claim 16, selected from among bacteria, yeasts, fungi, mosses, animal and plant organisms.

18. Cell cultures, parts, transgenic propagation material, or fruits derived from a transgenic organism according to Claims 16 or 17.
19. Use of a transgenic organism according to any one of Claims 16 or 17, or of cell cultures, parts, transgenic propagation material, or fruits derived therefrom according to Claim 18, as foodstuff or animal feed or for isolation of Vitamin E.
20. Antibodies, protein-binding, or DNA-binding factors against polypeptides with HGD, MAAI, or FAAH activity, genes or cDNAs thereof.
21. Use of polypeptides with HGD, MAAI, or FAAH activity, genes or cDNAs thereof to find inhibitors of HGD, MAAI, or FAAH.
22. A method for finding inhibitors of MAAI, HGD, or FAAH, wherein the enzyme activity of MAAI, HGD, or FAAH is measured in the presence of a chemical compound and if the enzyme activity decreases compared to the uninhibited activity, then the chemical compound constitutes an inhibitor.
23. Use of inhibitors of HGD, MAAI, or FAAH that can be obtained by a method according to Claim 22, as growth regulators.

[see German page for figure]

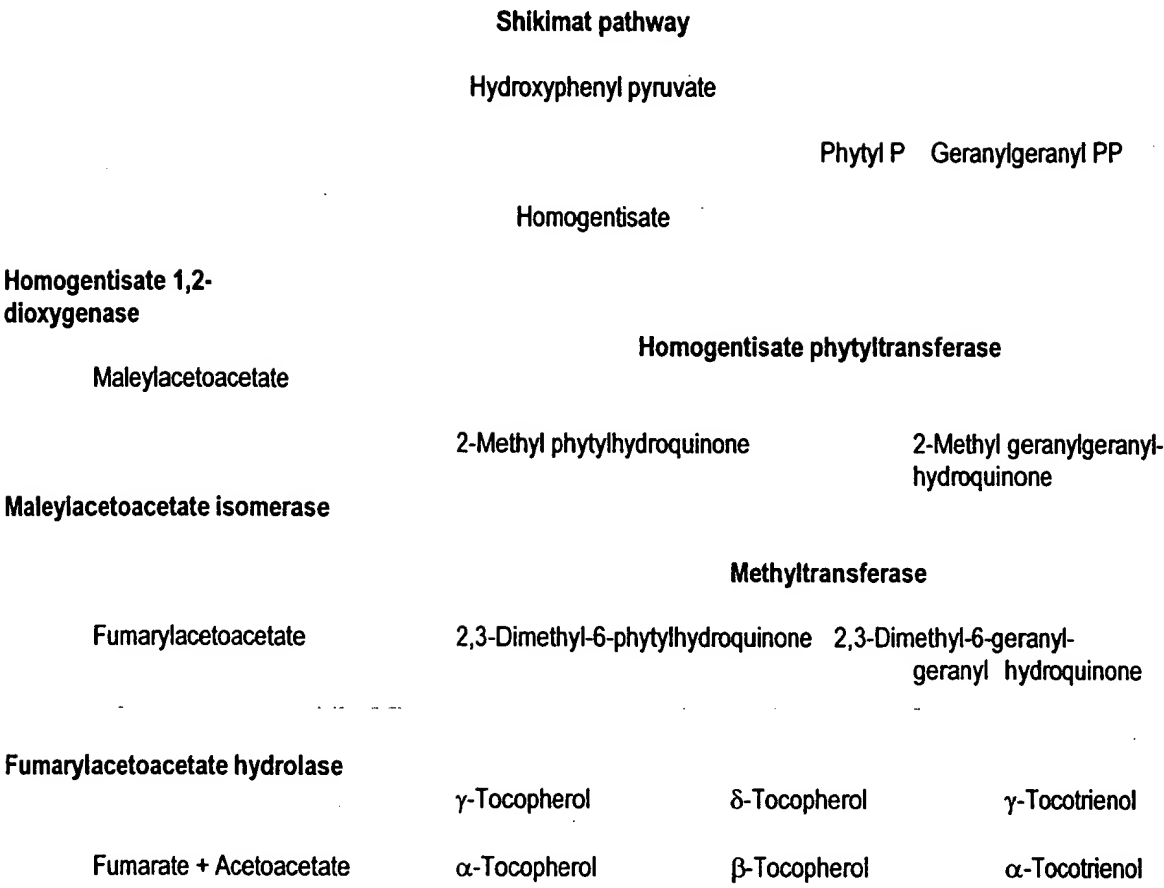


Figure 1

**[see German pages 2/9 through 9/9 for Figures 2-9, in English ]**

SEQUENCE PROTOCOL

<110> SunGene GmbH & Co. KGaA

<120> Improved methods for Vitamin E biosynthesis

[rest of this page in English]

**2-22**

[see pages 2-22, in English]

[German starts mid-page]

<213> Artificial sequence

<220>

<221> CDS

<222> (8) .. (1150)

<220>

<221> misc feature

<222> (1) .. (6)

<223> restriction site

<220>

<221> misc feature

<222> (1154) .. (1159)

<223> restriction site

<220>

<223> Description of artificial sequence: codon usage optimized cDNA coding for hydroxyphenylpyruvate dioxygenase from *Streptomyces avermitilis*

[rest of page in English]

[see page 24, in English]



[German starts near bottom of page]

<213> Artificial sequence

<223> Description of artificial sequence: codon usage optimized cDNA coding for hydroxyphenylpyruvate dioxygenase from *Streptomyces avermitilis*

[rest of page in English]

**26-44**

[see pages 26-44, in English]

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
18. April 2002 (18.04.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 02/31173 A2

(51) Internationale Patentklassifikation: C12P 17/06,  
C12N 15/11, 15/29, 15/52, 15/67, 15/74, 15/79, 1/21, 5/04,  
5/06, A23L 1/29, 1/302, A01H 5/00, 5/08, 13/00, 15/00

(74) Anwalt: DÖRPER, Thomas; Basf Aktiengesellschaft,  
67056 Ludwigshafen (DE).

(21) Internationales Aktenzeichen: PCT/EP01/10779

(22) Internationales Anmeldedatum:  
18. September 2001 (18.09.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
100 46 462.9 19. September 2000 (19.09.2000) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von  
US): SUNGENE GMBH & CO. KGaA [DE/DE]; 06468  
Gatersleben (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): GEIGER, Michael  
[DE/DE]; Neuer Weg 15, 06484 Quedlinburg (DE).  
EBNETH, Marcus [DE/DE]; Katzbachstr. 36, 10965  
Berlin (DE). KUNZE, Irene [DE/DE]; Mühlenweg 11,  
06466 Gatersleben (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,  
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE,  
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,  
LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,  
MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,  
YU, ZA, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW),  
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM), europäisches Patent (AT, BE, CH, CY, DE, DK,  
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR),  
OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu  
veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen  
Abkürzungen wird auf die Erklärungen ("Guidance Notes on  
Codes and Abbreviations") am Anfang jeder regulären Ausgabe  
der PCT-Gazette verwiesen.

(54) Title: IMPROVED METHOD FOR THE BIOSYNTHESIS OF VITAMIN E

(54) Bezeichnung: VERBESSERTE VERFAHREN ZUR VITAMIN E BIOSYNTHESE

(57) Abstract: The invention relates to improved methods for the biosynthesis of vitamin E. Said methods are characterized by the inhibition of the catabolization of homogentisate to maleylacetoacetate and fumarylacetoacetate and then to fumarate and acetoacetate. The invention also relates to the combination of this inhibition with methods that increase the provision of homogentisate or that promote the conversion of homogentisate to vitamin E. The invention further relates to nucleic acid constructs and vectors, which can be used to implement the inventive methods, in addition to transgenic plant organisms produced from said constructs and vectors.

(57) Zusammenfassung: Die Erfindung betrifft verbesserte Verfahren zur Biosynthese von Vitamin E. Diese Verfahren sind gekennzeichnet durch eine Inhibition des Abbaus von Homogentisat über Maleylacetoacetat und Fumarylacetoacetat zu Fumarat und Acetoacetat. Erfindungsgemäss ist ferner die Kombination dieser Inhibition mit Verfahren, die die Homogentisatbereitstellung steigern, bzw. die Umsetzung des Homogentisat zu Vitamin E fördern. Erfindungsgemäss sind Nukleinsäurekonstrukte und Vektoren, mit denen die erfindungsgemässen Verfahren realisiert werden können, sowie davon ausgehend erzeugte transgene pflanzliche Organismen.

WO 02/31173 A2

## Verbesserte Verfahren zur Vitamin E Biosynthese

### Beschreibung

5

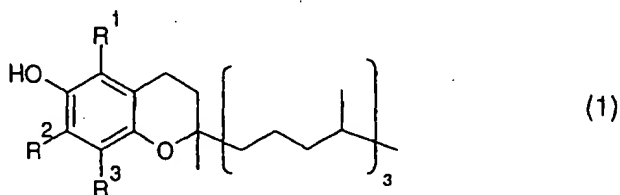
Die Erfindung betrifft verbesserte Verfahren zur Biosynthese von Vitamin E. Diese Verfahren sind gekennzeichnet durch eine Inhibition des Abbaus von Homogentisat (HG) über Maleylacetoacetat (MAA), Fumarylacetoacetat (FAA) zu Fumarat und Aceto-

10 acetat. Erfindungsgemäss ist ferner die Kombination dieser Inhibition mit Verfahren, die die Homogentisatbereitstellung weiter steigern, bzw. die Umsetzung des Homogentisat zu Vitamin E fördern.

15 Homogentisat ist ein bedeutender Stoffwechselmetabolit. Es ist ein Abbauprodukt der Aminosäuren Tyrosin und Phenylalanin. In Mensch und Tier wird Homogentisat weiter zu Maleylacetoacetat, infolge zu Fumarylacetoacetat und dann zu Fumarat und Acetoacetat abgebaut. Pflanzen und andere photosynthesebetreibende Mikro-  
20 organismen verwenden Homogentisat ferner als Ausgangsprodukt für die Synthese von Tocopherolen und Tocotrienolen.

Die in der Natur vorkommenden acht Verbindungen mit Vitamin E-Aktivität sind Derivate des 6-Chromanols (Ullmann's Encyclopedia of Industrial Chemistry, Vol. A 27 (1996), VCH Verlagsgesellschaft, Chapter 4., 478-488, Vitamin E). Die Gruppe der Tocopherole (1a-d) weist eine gesättigte Seitenkette auf, die Gruppe der Tocotrienole (2a-d) eine ungesättigte Seitenkette:

30



1a,  $\alpha$ -Tocopherol:  $R^1 = R^2 = R^3 = CH_3$

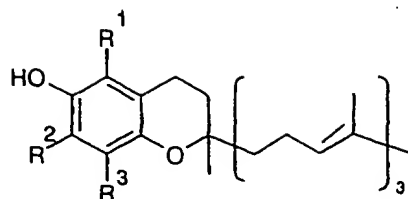
1b,  $\beta$ -Tocopherol [148-03-8]:  $R^1 = R^3 = \text{CH}_3$ ,  $R^2 = \text{H}$

40 1c,  $\gamma$ -Tocopherol [54-28-4]:  $R^1 = H$ ,  $R^2 = R^3 = CH_3$

1d,  $\delta$ -Tocopherol [119-13-1]:  $R^1 = R^2 = H$ ,  $R^3 = CH_3$

45

5



(2)

- 2a,  $\alpha$ -Tocotrienol [1721-51-3]:  $R^1 = R^2 = R^3 = \text{CH}_3$   
 10 2b,  $\beta$ -Tocotrienol [490-23-3]:  $R^1 = R^3 = \text{CH}_3$ ,  $R^2 = \text{H}$   
 2c,  $\gamma$ -Tocotrienol [14101-61-2]:  $R^1 = \text{H}$ ,  $R^2 = R^3 = \text{CH}_3$   
 2d,  $\delta$ -Tocotrienol [25612-59-3]:  $R^1 = R^2 = \text{H}$ ,  $R^3 = \text{CH}_3$

In der vorliegenden Erfindung werden unter Vitamin E alle acht  
 15 vorstehend erwähnten Tocopherole und Tocotrienole mit Vitamin-E-Aktivität verstanden.

Diese Verbindungen mit Vitamin-E-Aktivität sind wichtige natürliche fettlösliche Antioxidantien. Ein Mangel an Vitamin E führt  
 20 bei Menschen und Tieren zu pathophysiologischen Situationen. Epidemiologische Untersuchungen haben ergeben, dass eine Nahrungsergänzung mit Vitamin E das Risiko, an kardiovaskuläre Erkrankungen oder Krebs zu erkranken, reduziert. Ferner ist eine positive Wirkung auf das Immunsystem und eine Prävention von  
 25 generellen altersbedingten Degenerationerscheinungen beschrieben (Traber MG, Sies H; Annu Rev Nutr. 1996;16:321-47). Die Funktion von Vitamin E liegt dabei vermutlich in einer Stabilisierung der biologischen Membranen, sowie in einer Reduktion von freien Radikalen, wie sie zum Beispiel bei der Lipidoxidation von poly-  
 30 ungesättigten Fettsäuren (PUFA) entstehen.

Die Funktion von Vitamin E in den Pflanzen selber ist kaum untersucht. Es scheint jedoch eventuell eine wichtige Funktion in der Stressreaktion der Pflanze v.a. auf oxidativen Stress zu haben.  
 35 Erhöhte Vitamin E Spiegel wurden mit erhöhter Stabilität und Lagerfähigkeit von Pflanzenprodukten in Verbindung gebracht. Der Zusatz von Vitamin E zu Tierernährungsprodukten hat einen positiven Effekt auf die Fleischqualität und die Lagerfähigkeit von Fleisch und Fleischprodukten bei z.B. Schweinen, Rindern und  
 40 Geflügel.

Vitamin E-Verbindungen haben daher einen hohen wirtschaftlichen Wert als Zusatzstoffe im Food- und Feed-Bereich, in pharmazeutischen Formulierungen und in kosmetischen Anwendungen.

## 3

In der Natur wird Vitamin E ausschliesslich von Pflanzen und anderen photosynthetisch aktiven Organismen (z.B. Cyanobakterien) synthetisiert. Der Gehalt an Vitamin E variiert stark. Die meisten Grundnahrungsmittelpflanzen (z.B. Weizen, Reis, Mais, 5 Kartoffel) haben einen nur sehr geringen Vitamin E Gehalt (Hess, Vitamin E,  $\alpha$ -tocopherol, In *Antioxidants in Higher Plants*, Herausgeber: R. Ascher and J. Hess, 1993, CRC Press, Boca Raton, pp. 111-134). Ölsaaten haben in der Regel einen deutlich höher Vitamin E Gehalt, wobei  $\beta$ -,  $\gamma$ -,  $\delta$ -Tocopherol dominieren. Die 10 empfohlene Tagesdosis an Vitamin E liegt bei 15-30mg.

Fig. 1 zeigt ein Biosyntheseschema von Tocopherolen und Tocotrienolen.

15 Im Verlauf der Biosynthese wird Homogentisinsäure (Homogentisat; HG) an Phytylpyrophosphat (PPP) bzw. Geranylgeranylpyrophosphat gebunden, um die Vorläufer von  $\alpha$ -Tocopherol und  $\alpha$ -Tocotrienol, das 2-Methyl-phytylhydrochinon bzw. das 2-Methyl-geranylgeranylhydrochinon zu bilden. Durch Methylierungsschritte mit S-Adenosylmethionin als Methyl-Gruppen-Donor entsteht zunächst 2,3-Dimethyl-6-phytylhydrochinon, dann durch Zyklisierung  $\gamma$ -Tocopherol und durch nochmalige Methylierung  $\alpha$ -Tocopherol. Ferner kann aus 2-Methyl-phytylhydrochinon durch Methylierung  $\beta$ - und  $\delta$ -Tocopherol synthetisiert werden.

25

Über die Erhöhung des Stoffwechselflusses zur Steigerung des Tocopherol- bzw. Tocotrienolgehaltes in transgenen Organismen, beispielsweise in transgenen Pflanzen durch Überexpression einzelner Biosynthesegene ist bisher wenig bekannt.

30

WO 97/27285 beschreibt eine Modifikation des Tocopherol-Gehaltes durch verstärkte Expression bzw. durch Herunterregulation des Enzyms p-Hydroxyphenyl-pyruvatdioxygenase (HPPD).

35 WO 99/04622 beschreibt Gensequenzen codierend für eine  $\gamma$ -Tocopherol-methyltransferase aus *Synechocystis* PCC6803 und *Arabidopsis thaliana* und deren Einbau in transgene Pflanzen.

WO 99/23231 zeigt, dass die Expression einer Geranylgeranyloxidoreduktase in transgenen Pflanzen eine gesteigerte Tocopherolbiosynthese zur Folge hat. 40

WO 00/10380 zeigt eine Veränderung der Zusammensetzung an Vitamin E unter Verwendung von 2-Methyl-6-phytylplastoquinol-methyltransferase. 45

Shintani and DellaPenna haben gezeigt, dass eine Überexpression der  $\gamma$ -Tocopherolmethyltransferase die Vitamin E Gehalt deutlich steigern kann (Shintani und Dellapenna, Science 282 (5396):2098-2100, 1998).

5

Alle Reaktionen der Vitamin E Biosynthese laufen über das Homogentisat. Bisherige Untersuchungen haben sich meist auf die Überexpression von Genen der Vitamin E- oder Homogentisat-Biosynthese beschränkt (siehe oben). Wenig Beachtung wurde bislang den Konkurrenzreaktionen zuteil, die Homogentisat abbauen und so der Vitamin E Biosynthese entziehen.

Der Abbau von Homogentisat über Maleylacetoacetat und Fumarylacetoacetat zu Fumarat und Acetoacetat ist für nicht photosynthetisch aktive Organismen, vor allem tierische Organismen, beschrieben (Fernandez-Canon JM et al., Proc Natl Acad Sci USA. 1995; 92 (20):9132-9136). Tierische Organismen benutzen diesen Stoffwechselweg zum Abbau aromatischer Aminosäuren, die vorwiegend über die Nahrung aufgenommen werden. Seine Funktion und Relevanz in Pflanzen ist hingegen unklar. Die Reaktionen werden durch die Homogentisat-1,2-dioxygenase (HGD; EC-Nr.: 1.13.11.5), die Maleylacetoacetatisomerase (MAAI; EC-Nr.: 5.2.1.2.) und die Fumarylacetoacetathydrolase (FAAH; EC-Nr.: 3.7.1.2) katalysiert.

Das Gen der Homogentisat-1,2-dioxygenase (HGD) aus *Arabidopsis thaliana* ist bekannt (Genbank Acc.-No. AF130845). Das Gen der Fumarylacetoacetathydrolase aus *Arabidopsis thaliana* war bereits aufgrund einer Homologie zu der Fumarylacetoacetathydrolase aus *Emmericella nidulans* (gb|L41670) als ähnlich zu derselben annotiert (Genbank Acc.-No. AC002131). Es ist jedoch in dem entsprechenden Genbank-Eintrag ausdrücklich darauf hingewiesen, dass die Annotation allein auf Ähnlichkeit und nicht auf experimentellen Daten beruht. Das Gen der Maleylacetoacetatisomerase (MAAI) aus *Arabidopsis* war als Gen in der Genbank vorhanden (AC005312), jedoch als eine putative Glutathion-S-Transferase annotiert. Bekannt war eine MAAI aus *Emmericella nidulans* (Genbank Acc.-No. EN 1837).

Tsegaye et al. mutmassen in einem Abstract (Abstract No. 413) zum 1999 Jahrestreffen der American Society of Plant Physiologists (24.-28.07.1999, Baltimore, USA) einen Vorteil in der Kombination einer Kreuzung von HPPD-überexprimierenden Pflanzen mit Pflanzen in denen die HGD durch einen antisense-Ansatz herunterreguliert wird.

45

Trotz einiger Erfolge besteht weiterhin Bedarf an einer Optimierung der Vitamin E Biosynthese. Der Erfindung lag die Aufgabe zugrunde, weitere Verfahren zur Verfügung zu stellen, die den Vitamin E-Biosyntheseweg beeinflussen und damit zu weiter vor-  
5 teilhaften transgenen Pflanzen mit erhöhtem Gehalt an Vitamin E führen.

Die Aufgabe wurde durch Identifikation des Homgentisat-Maleyl-acetoacetat-Fumarylacetoacetat-Fumarat Abbauweges als wesent-  
10 licher Konkurrenzweg zu dem Vitamin E-Biosyntheseweg gelöst. Es wurde gefunden, dass eine Inhibition dieses Abbauweges zu einer Optimierung der Vitamin E-Biosynthese führt.

Ein erster Gegenstand der Erfindung betrifft daher Verfahren zu  
15 einer Vitamin E Produktion durch Reduktion der HGD, MAAI bzw. FAAH Aktivität. Als besonders vorteilhaft erweist sich eine Kombination der beschriebenen Inhibition des Homogentisat-Abbauweges mit anderen Verfahren, die zu einer verbesserten Vitamin E Biosynthese führen, indem sie die Umsetzung vom Homogentisat zu  
20 Vitamin E fördern. Dies kann durch eine erhöhte Zuverfügungstellung von Reaktionspartnern oder durch eine erhöhte Umsetzung des Homogentisates mit eben diesen realisiert werden. Beispielfürhaft lässt sich dieser Effekt mit einer Überexpression der Homogentisatphytyltransferase (HGPT), Geranylgeranyloxidoreduktase,  
25 der 2-Methyl-6-phytylplastoquinol-methyltransferase oder der  $\gamma$ -Tocopherol-methyltransferase erreichen.

Vorteilhaft ist ferner eine Kombination mit Genen die die Homogentisatbildung fördern, wie zum Beispiel der HPPD oder dem TyrA-  
30 Gen.

Die Inhibition des Abbauweges vom Homogentisat, über Maleyl-acetoacetat und Fumarylacetoacetat zum Fumarat und Acetoacetat kann auf vielfältige Art und Weise realisiert werden.

35 Ein Gegenstand der Erfindung betrifft Nukleinsäurekonstrukte, die wenigstens eine Nukleinsäuresequenz (anti-MAAI/FAAH), welche zu einer Inhibition des Maleylacetoacetat-Fumarylacetoacetat-Fumarat Weges befähigt ist, oder ein funktionales Äquivalent davon ent-  
40 halten.

Ein weiterer Gegenstand der Erfindung betrifft oben beschriebene Nukleinsäurekonstrukte, die neben der anti-MAAI/FAAH Nukleinsäuresequenz zusätzlich wenigstens eine Nukleinsäuresequenz  
45 (pro-HG), welche zu einer Steigerung der Biosynthese von Homogentisat (HG) befähigt ist, oder ein funktionales Äquivalent davon oder wenigstens eine Nukleinsäuresequenz (pro-VitaminE), welche



6

zu einer Steigerung der Vitamin E Biosynthese ausgehend vom Homogentisat befähigt ist, oder ein funktionales Äquivalent davon oder eine Kombination von pro-HG und pro-VitaminE bzw. ihrer funktionellen Äquivalente enthalten.

5

Ein weiterer Gegenstand der Erfindung betrifft Nukleinsäurekonstrukte, die eine Nukleinsäuresequenz (anti-HGD) enthalten, welche zu einer Inhibition der Homogentisat-1,2-dioxygenase (HGD) befähigt ist, oder für ein funktionales Äquivalent davon.

10

Ferner betrifft die Erfindung besagte anti-HGD Nukleinsäurekonstrukte, die neben der anti-HGD Nukleinsäuresequenz zusätzlich wenigstens eine Nukleinsäuresequenz, die für eine bifunktionale Chorismatmutase-Prephenatdehydrogenase (TyrA) oder deren funktionelle Äquivalente, oder wenigstens eine Nukleinsäuresequenz (pro-Vitamin E), welche zu einer Steigerung der Vitamin E Biosynthese ausgehend vom Homogentisat befähigt ist, oder deren funktionale Äquivalente, bzw. eine Kombination von pro-VitaminE und TyrA-Sequenzen oder deren funktionellen Äquivalenten enthalten.

20

TyrA kodiert für ein bifunktionale Chorismatmutase-Prephenatdehydrogenase aus E.coli, eine Hydroxyphenylpyruvatsynthase, dass die enzymatischen Aktivitäten einer Chorismatmutase und Prephenatdehydrogenase beinhaltet, und Chorismat zu Hydroxyphenylpyruvat, dem Homogentisatedukt, umsetzt (Christendat D, Turnbull JL. Biochemistry. 1999 Apr 13;38(15):4782-93; Christopherson RI, Heyde E, Morrison JF. Biochemistry. 1983 Mar 29;22(7):1650-6.).

25

Ferner betrifft die Erfindung Nukleinsäurekonstrukte, die wenigstens eine Nukleinsäuresequenz (pro-HG), welche zu einer Steigerung der Homogentisat (HG) Biosynthese befähigt ist, oder ein funktionales Äquivalent davon, und wenigstens eine Nukleinsäuresequenz (pro-Vitamin E), welche zu einer Steigerung der Vitamin E Biosynthese ausgehend vom Homogentisat befähigt ist, oder ein funktionales Äquivalent davon, enthalten.

35

Erfindungsgemäss sind weiterhin funktionelle Analoga der oben erwähnten Nukleinsäurekonstrukte. Funktionelle Analoga meint hier zum Beispiel eine Kombination der einzelnen Nukleinsäuresequenzen

40

1. auf einem Polynukleotid (Mehrfachkonstrukte)
2. auf mehreren Polynukleotiden in einer Zelle (Kotransformation)

45

3. durch Kreuzung verschiedener transgener Pflanzen, die jeweils mindestens eine der besagten Nukleotidsequenzen enthalten.

Bevorzugt sind die in den Nukleinsäurekonstrukte enthaltenen Nukleinsäuresequenzen funktionell mit genetischen Kontrollsequenzen verbunden.

Die erfindungsgemässe Transformation von Pflanzen mit einem pro-HG kodierenden Konstrukt führt zur Steigerung der Homogentisatbildung. Durch gleichzeitige Transformation mit anti-HGD bzw. anti-MAAI/FAAH, insbesondere dem anti-MAAI Konstrukt, wird ein unerwünschter Abfluss dieses Metaboliten vermieden. Eine erhöhte Homogentisatmenge steht in der transgenen Pflanze somit zur Bildung von Vitamin E, zum Beispiel von Tocopherolen, über die Intermediate Methyl-6-phytylquinol und 2,3-Dimethyl-phytylquinol (vgl. Fig. 1), zur Verfügung. Sowohl pro-HG als auch anti-MAAI/FAAH oder anti-HGD führen zu einer erhöhten Homogentisatbereitstellung zur Vitamin E-Biosynthese. Die Umsetzung von Homogentisat zu Vitamin E kann durch eine kombinierte Transformation mit einem pro-VitaminE kodierenden Konstrukt verbessert werden und erhöht weiterhin die Biosynthese von Vitamin E.

"Steigerung" der Homogentisat-Biosynthese ist in diesem Zusammenhang weit auszulegen und umfasst die Erhöhung der Homogentisat (HG)-Biosyntheseaktivität in der mit einem erfindungsgemässen pro-HG Konstrukt transformierten Pflanze oder dem Pflanzenteil oder Gewebe. Erfindungsgemäss sind verschiedene Strategien zur Erhöhung der HG-Biosyntheseaktivität umfasst. Der Fachmann erkennt, dass eine Reihe verschiedener Methoden zur Verfügung steht, um die HG-Biosyntheseaktivität in gewünschter Weise zu beeinflussen. Die infolge beschriebenen Verfahren sind insofern beispielhaft und nicht einschränkend zu verstehen.

Die erfindungsgemäss bevorzugte Strategie umfasst die Verwendung einer Nukleinsäuresequenz (pro-HG), die transkribiert und zu einem Polypeptid translatiert werden kann, das die HG-Biosyntheseaktivität steigert. Beispiele für derartige Nukleinsäuresequenzen sind die p-Hydroxyphenylpyruvatoxygenase (HPPD) aus verschiedenen Organismen oder das bakterielle TyrA-Genprodukt. Neben der beschriebenen künstlichen Expression von bekannten Genen, kann man auch deren Aktivität durch Mutagenese der Polypeptidsequenz erhöhen. Ferner ist eine gesteigerte Transkription und Translation der endogenen Gene zum Beispiel durch Verwendung künstlicher Transkriptionsfaktoren vom Typ der Zinkfingerproteine zu erreichen (Beerli RR et al., Proc Natl Acad Sci U S A. 2000; 97 (4):1495-500). Diese Faktoren lagern sich in der regulatorischen Bereichen der endogenen Gene an und bewirken, je nach Ge-

staltung des Faktors, eine Expression oder Repression des endogenen Gens.

Besonders bevorzugt für pro-HG ist die Verwendung der durch  
5 Nukleinsäuren, die für Polypeptide gemäss SEQ ID NO: 8, 11 oder 16 kodieren, besonders bevorzugt Nukleinsäuren mit den durch SEQ ID NO: 7, 10 oder 15 beschriebenen Sequenzen.

Analog ist die "Steigerung" der Vitamin E Biosyntheseaktivität  
10 zu verstehen, wobei hier Gene zum Einsatz kommen, deren Aktivität die Umsetzung von Homogentisat zu Vitamin E (Tocopherolen, Tocotrienolen) oder die Synthese von Reaktionspartnern des Homogentisats, wie zum Beispiel des Phytylpyrophosphat oder Geranylgeranylpyrophosphat, fördert. Beispielfhaft seien genannt die  
15 Homogentisat-phytyltransferase (HGPT), Geranylgeranyloxidoreduktase, 2-Methyl-6-phytylplastoquinol-methyltransferase und  $\gamma$ -Tocopherol-methyltransferase. Besonders bevorzugt ist die Verwendung von Nukleinsäuren, die für Polypeptide gemäss SEQ ID NO: 14, 20, 22 oder 24 kodieren, besonders bevorzugt  
20 sind mit den durch SEQ ID NO: 13, 19, 21 oder 23 beschriebenen Sequenzen.

"Inhibition" ist im Zusammenhang mit anti-MAAI/FAAH bzw. anti-HGD weit auszulegen und umfasst die teilweise oder im wesentlichen  
25 vollständige, auf unterschiedliche zellbiologische Mechanismen beruhende Unterbindung oder Blockierung der MAAI/FAAH- bzw. HGD-Enzymaktivität in der mit einem erfindungsgemässen anti-MAAI/FAAH- bzw. anti-HGD-Konstrukt transformierten Pflanze oder dem Pflanzenteil oder Gewebe. Eine Inhibition im Sinne der Erfindung  
30 umfasst auch eine mengenmässige Verringerung von aktiver HGD, MAAI oder FAAH in der Pflanze, bis hin zu einem im wesentlichen vollständigen Fehlen von HGD, MAAI oder FAAH-Protein (d.h. fehlende Nachweisbarkeit von HGD bzw. MAAI oder FAAH-Enzymaktivität oder fehlende immunologische Nachweisbarkeit von HGD, MAAI oder  
35 FAAH).

Erfindungsgemäss sind verschiedene Strategien zur Verringerung oder Inhibition der HGD bzw. MAAI oder FAAH-Aktivität umfasst. Der Fachmann erkennt, dass eine Reihe verschiedener Methoden zur  
40 Verfügung steht, um die HGD bzw. MAAI oder FAAH-Genexpression oder Enzymaktivität in gewünschter Weise zu beeinflussen.

Die erfindungsgemäss bevorzugte Strategie umfasst die Verwendung einer Nukleinsäuresequenz (anti-MAAI/FAAH bzw. anti-HGD), welche  
45 zu einer antisense-Nukleinsäuresequenz transkribierbar ist, die zur Inhibition der HGD- bzw. MAAI/FAAH-Aktivität befähigt ist,

z.B. indem sie die Expression von endogener HGD bzw. MAAI oder FAAH inhibiert.

Die erfindungsgemässen anti-HGD bzw. anti-MAAI/FAAH-Nukleinsäure-  
5 sequenzen können gemäss einer bevorzugten Ausführungsform die in antisense-Orientierung insertierte kodierende Nukleinsäuresequenz der HGD (anti-HGD) bzw. MAAI oder FAAH (anti-MAAI/FAAH) oder funktional äquivalente Fragment der jeweiligen Sequenzen enthalten.

10

Besonders bevorzugte anti-HGD-Nukleinsäuresequenzen umfassen Nukleinsäuresequenzen, welche für Polypeptide enthaltend eine Aminosäuresequenz gemäss SEQ ID NO: 3 oder funktionelle Äquivalente davon kodieren. Besonders bevorzugt sind Nukleinsäure-  
15 sequenzen gemäss SEQ ID NO: 1, 2 oder 12 oder funktionelle Äquivalente davon.

Besonders bevorzugte anti-MAAI/FAAH-Nukleinsäuresequenzen umfassen Nukleinsäuresequenzen, welche für Polypeptide ent-  
20 haltend eine Aminosäuresequenz gemäss SEQ ID NO: 5 und 18 oder funktionelle Äquivalente davon kodieren. Besonders bevorzugt sind Nukleinsäuresequenzen gemäss SEQ ID NO: 4, 6, 9 oder 17 oder funktionelle Äquivalente davon, ganz besonders bevorzugt sind die mit SEQ ID NO: 41 oder 42 wiedergegebenen Teilsequenzen oder  
25 deren funktionelle Äquivalente.

Eine bevorzugte Ausführungsform der erfindungsgemässen Nukleinsäuresequenzen umfasst ein HGD-, MAAI- oder FAAH- Sequenzmotiv gemäss SEQ ID NO: 1, 2, 4, 6, 9, 12, 17, 41 oder 42 in antisense-  
30 Orientierung. Dies führt zur vermehrten Transkription von Nukleinsäuresequenzen in der transgenen Pflanze, welche komplementär zur endogenen kodierenden HGD, MAAI bzw. FAAH-Sequenz oder einem Teil davon sind und mit dieser auf DNA- oder RNA-Ebene hybridisieren.

35

Vorteilhaft kann die antisense-Strategie mit einem Ribozym-Verfahren gekoppelt werden. Ribozyme sind katalytisch aktive RNA Sequenzen, die gekoppelt an die antisense Sequenzen, die Zielsequenzen katalytisch spalten (Tanner NK. FEMS Microbiol Rev. 1999;  
40 23 (3):257-75). Dies kann die Effizienz einer anti-sense Strategie erhöhen.

Weitere Methoden zur Inhibition der HGD- bzw. MAAI/FAAH-Expression umfassen die zu Kosuppression führende Überexpression  
45 homologer HGD- bzw. MAAI/FAAH-Nukleinsäuresequenzen (Jorgensen et al., Plant Mol. Biol. 1996, 31 (5):957-973), die Induktion des spezifischen RNA-Abbaus durch die Pflanze mit Hilfe eines viralen

Expressionssystem (Amplikon) (Angell, SM et al., Plant J. 1999, 20(3):357-362). Diese Methoden werden auch als "post-transcriptional gene silencing" (PTGS) bezeichnet.

- 5 Weitere Methoden sind die Einführung von Nonsense-Mutationen in das Endogen mittels Einführung von RNA/DNA-Oligonukleotiden in die Pflanze (Zhu et al., Nat. Biotechnol. 2000, 18(5):555-558) oder die Generierung von Knockout-Mutanten mit Hilfe von z.B. T-DNA-Mutagenese (Koncz et al., Plant Mol. Biol. 1992, 20(5):963-976) oder homologer Rekombination (Hohn, B. und Puchta, H, Proc. Natl. Acad. Sci. USA. 1999, 96:8321-8323.). Ferner ist eine Genüberexpression oder -repression auch mit spezifischen DNA-bindenden Faktoren z.B. mit den oben erwähnten Faktoren vom Typus der Zinkfingertranskriptionsfaktoren. Ferner können Fakto-
- 10 ren in eine Zelle eingebracht werden, die das Zielprotein selber inhibieren. Die Protein-bindenden Faktoren können z.B. Aptamere sein (Famulok M, und Mayer G. Curr Top Microbiol Immunol. 1999; 243:123-36).
- 20 Auf die oben beschriebenen Druckschriften und die darin offenbarten Methoden zur Regulation der pflanzlichen Genexpression wird hiermit ausdrücklich Bezug genommen.

Eine anti-HGD bzw. anti-MAAI/FAAH Sequenz im Sinne der vor-

25 liegenden Erfindung ist somit insbesondere ausgewählt unter:

- a) antisense-Nukleinsäuresequenzen;
- b) antisense-Nukleinsäuresequenzen kombiniert mit einem
- 30 Ribozym-Verfahren
- c) für homologe HGD- bzw. MAAI/FAAH-kodierende und zu Kosuppression führende Nukleinsäuresequenzen
- 35 d) HGD- bzw. MAAI/FAAH-RNA-Abbau bewirkende virale Nukleinsäuresequenzen und Expressionskonstrukte;
- e) Nonsense-Mutanten von endogenen HGD- bzw. MAAI/FAAH kodierenden Nukleinsäuresequenzen;
- 40 f) für Knockout-Mutanten kodierende Nukleinsäuresequenzen;
- g) zu homologer Rekombination geeignete Nukleinsäuresequenzen;

- h) Nukleinsäuresequenzen kodierend für spezifische DNA- oder Protein-bindende Faktoren mit anti-HGD- bzw. anti-MAAI/FAAH-Aktivität

- 5 wobei die Expression jeder einzelner dieser anti-HGD bzw. anti-MAAI/FAAH Sequenzen eine "Inhibition" der HGD- bzw. MAAI/FAAH-Aktivität im Sinne der Erfindung bewirken kann. Auch eine kombinierte Anwendung solcher Sequenzen ist denkbar.
- 10 Unter Nukleinsäurekonstrukt oder Nukleinsäuresequenz versteht man erfindungsgemäss beispielsweise eine genomische oder eine komplementäre DNA-Sequenz oder eine RNA-Sequenz sowie halb- oder vollsynthetische Analoga davon. Diese Sequenzen können in linearer oder zirkulären Form, extrachromosomal oder integriert in das
- 15 Genom vorliegen. Die pro-HG, pro-VitaminE, anti-HGD bzw. anti-MAAI/FAAH Nukleotidsequenzen der erfindungsgemässen Konstrukte können synthetisch hergestellt oder natürlich gewonnen werden oder eine Mischung aus synthetischen und natürlichen DNA-Bestandteilen enthalten, sowie aus verschiedenen heterologen HGD, MAAI/
- 20 FAAH, pro-HG oder pro-VitaminE Genabschnitten verschiedener Organismen bestehen. Die anti-HGD bzw. anti-MAAI/FAAH-Sequenz kann von einem oder mehreren Exons oder Introns, insbesondere Exons der HGD, MAAI oder FAAH-Gene abgeleitet sein.
- 25 Ausserdem sind artifizielle Nukleinsäuresequenzen geeignet, solange sie, wie oben beschrieben, die gewünschte Eigenschaft beispielsweise der Erhöhung des VitaminE-Gehaltes in der Pflanze durch Überexpression mindestens eines proHG- und/oder proVitamin E-Gens und/oder Expression einer anti-HGD bzw. MAAI/FAAH-Sequenz
- 30 in Kulturpflanzen vermitteln. Beispielsweise können synthetische Nukleotid-Sequenzen mit Kodons erzeugt werden, die von den zu transformierenden Pflanzen bevorzugt werden. Diese von Pflanzen bevorzugten Kodons können anhand der Kodonnutzung aus Kodons mit der höchsten Proteinhäufigkeit in üblicher Weise bestimmt werden.
- 35 Solche artifiziellen Nukleotid-Sequenzen können beispielsweise durch Rückübersetzung mittels Molecular Modelling konstruierter Proteine, die HGD bzw. MAAI/FAAH- bzw. proHG-Aktivität oder proVitamin E-Aktivität aufweisen oder durch in vitro-Selektion ermittelt werden. Besonders geeignet sind kodierende Nukleotid-
- 40 Sequenzen, die durch Rückübersetzung einer Polypeptidsequenz gemäss der für die Wirtspflanze spezifischen Kodon-Nutzung erhalten wurden. Um unerwünschte pflanzliche Regulationsmechanismen zu umgehen, kann man beispielsweise ausgehend von der Aminosäuresequenz einer bakteriellen pro-HG, zum Beispiel des bakteriellen
- 45 TyrA Gens, und unter Berücksichtigung der pflanzlichen Kodon-Nutzung DNA-Fragmente rückübersetzen und daraus die vollständige, für einen Einsatz in der Pflanze optimierte exogene proHG-Sequenz

herstellen. Daraus wird ein proHG-Enzym exprimiert, welches der pflanzlichen Regulation nicht oder nur unzureichend zugänglich ist, wodurch die Überexpression von Enzymaktivität voll zur Geltung gelangen kann.

5

- Alle vorstehend erwähnten Nukleotid-Sequenzen sind in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Bei der Präparation eines Nukleinsäurekonstruktes können verschiedene DNA-Fragmente so manipuliert werden, dass eine Nukleotid-Sequenz mit korrekter Leserichtung und korrektem Leseraster erhalten wird. Für die Verbindung der Nukleinsäure-Fragmente untereinander können an die Fragmente Adaptoren oder Linker angesetzt werden. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mit Hilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.
- 25 Funktionale Äquivalente der pro-HG oder pro-VitaminE Sequenzen sind solche Sequenzen, welche trotz abweichender Nukleotidsequenz noch für ein Protein mit den erfindungsgemäss gewünschten Funktionen kodieren, d.h für ein Enzym mit direkt oder indirekt die Homogentisatbildung steigernder Aktivität (pro-HG), bzw. für ein Enzym mit direkt oder indirekt die Umsetzung von Homogentisat zu Vitamin E fördernder Aktivität (pro-Vitamin E).
- 30

- Funktionale Äquivalente von anti-HGD bzw. anti-MAAI/FAAH umfassen solche Nukleotidsequenzen welche die HGD bzw. MAAI/FAAH-Enzymfunktion in der transgenen Pflanze in ausreichendem Masse unterbinden. Dies kann z.B. durch Behinderung oder Unterbindung der HGD bzw. MAAI/FAAH-Prozessierung, des Transports von HGD bzw. MAAI/FAAH oder deren mRNA, Hemmung der Ribosomenanlagerung, Hemmung des RNA-Spleissens, Induktion eines RNA-abbauenden Enzyms und/oder Hemmung der Translationselongation oder -termination erfolgen. Ferner ist eine direkte Repression der endogenen Gene durch DNA-bindende Faktoren, zum Beispiel vom Typ der Zinkfinger-Transkriptionsfaktoren, möglich. Auch eine direkte Inhibition der entsprechenden Polypeptide, zum Beispiel durch Aptamere, ist machbar. Verschiedene Beispiele sind oben gegeben.
- 45

Unter einem funktionalen Äquivalent versteht man insbesondere auch natürliche oder künstliche Mutationen einer ursprünglich isolierten für HGD bzw. MAAI/FAAH oder pro-HG oder pro-Vitamin E kodierenden Sequenz, welche weiterhin die gewünschte Funktion zeigen. Mutationen umfassen Substitutionen, Additionen, Deletionen, Vertauschungen oder Insertionen eines oder mehrerer Nukleotidreste. Somit werden beispielsweise auch solche Nukleotidsequenzen durch die vorliegende Erfindung mit umfasst, welche man durch Modifikation der HGD bzw. MAAI/FAAH- bzw. proHG oder proVitamin E-Nukleotidsequenz erhält. Ziel einer solchen Modifikation kann z.B. die weitere Eingrenzung der darin enthaltenen kodierenden Sequenz oder z.B. auch die Einfügung weiterer Restriktionsenzym-Schnittstellen oder die Entfernung überflüssiger DNA sein.

15 Wo Insertionen, Deletionen oder Substitutionen, wie z.B. Transitionen und Transversionen, in Frage kommen, können an sich bekannte Techniken, wie in vitro-Mutagenese, "primer repair", Restriktion oder Ligation verwendet werden. Durch Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Überhängen für "blunt ends" können komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden.

Unter Substitution ist der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sogenannte konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähnliche Eigenschaft hat wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

30 Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptides und die Verknüpfungen zwischen den einzelnen Proteindomänen.

35 Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren ersetzt wird.

40 Unter Homologie zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinnlänge verstanden, die durch Vergleich mit Hilfe des Programmalgorithmus GAP (UWGCG, University of Wisconsin, Genetic Computer Group) unter Einstellung folgender Parameter berechnet wird:



14

Gap Weight:	12	Length Weight:	4
Average Match:	2,912	Average Mismatch:	-2,003

Unter einer Sequenz, die eine Homologie von mindestens 20 %  
5 auf Nukleinsäurebasis mit der Sequenz SEQ ID NO. 6 aufweist,  
wird dementsprechend eine Sequenz verstanden, das bei einem Ver-  
gleich seiner Sequenz mit der Sequenz SEQ ID NO. 6 nach obigem  
Programmalgorithmus mit obigem Parametersatz eine Homologie von  
mindestens 20 % aufweist.

10

Funktionelle Äquivalente, abgeleitet von einer der in den erfin-  
dungsgemässen Nukleinsäurekonstrukten oder Vektoren zum Einsatz  
kommenden Nukleinsäuresequenzen zum Beispiel durch Substitution,  
Insertion oder Deletion von Aminosäuren bzw. Nukleotiden, haben  
15 eine Homologie von mindestens 20 %, bevorzugt 40 %, vorzugsweise  
mindestens 60 %, bevorzugt mindestens 80 %, besonders bevorzugt  
mindestens 90 %.

Weitere Beispiele für die in den erfindungsgemässen Nukleinsäure-  
20 konstrukten oder Vektoren zum Einsatz kommenden Nukleinsäurese-  
quenzen lassen sich beispielsweise aus verschiedenen Organismen,  
deren genomische Sequenz bekannt ist, wie beispielsweise aus Ara-  
bidopsis thaliana durch Homologievergleiche der Aminosäuresequen-  
zen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen  
25 aus Datenbanken leicht auffinden.

Funktionale Äquivalente umfassen auch solche Varianten, deren  
Funktion, verglichen mit dem Ausgangsgen bzw. Genfragment, abge-  
schwächt oder verstärkt ist, also beispielsweise solche proHG-  
30 oder proVitamin E-Gene, welche für eine Polypeptid-Variante mit  
niedrigerer oder höherer enzymatischer Aktivität als der des  
Ursprungsgens kodieren.

Als weitere geeignete funktionell äquivalente Nukleinsäuresequen-  
35 zen sind Sequenzen zu nennen, welche für Fusionsproteine kodie-  
ren, wobei Bestandteil des Fusionsproteins z.B. ein proHG- bzw.  
proVitamin E-Polypeptid oder ein funktionell äquivalenter Teil  
davon ist. Der zweite Teil des Fusionsproteins kann z.B. ein wei-  
teres Polypeptid mit enzymatischer Aktivität (zum Beispiel ein  
40 weiteres proHG- bzw. proVitaminE-Polypeptid oder ein funktionell  
äquivalenter Teil davon) sein oder eine antigene Polypeptid-  
sequenz, mit deren Hilfe ein Nachweis der proHG- oder proVita-  
min E-Expression möglich ist (z.B. myc-tag oder his-tag). Bevor-  
zugt handelt es sich dabei jedoch um eine regulative Proteinse-  
45 quenz, wie z.B. ein Signal- oder Transitpeptid, das das proHG-  
oder proVitamin E-Protein an den gewünschten Wirkort leitet.

Ein weiterer Gegenstand der Erfindung betrifft rekombinante Vektoren, die wenigstens ein Nukleinsäurekonstrukt gemäss obiger Definition, eine Nukleinsäuresequenz, die für eine HGD, MAAI oder FAAH kodiert, oder Kombinationen dieser Möglichkeiten umfassen.

5

Bevorzugt sind die in den Vektoren enthaltenen Nukleinsäuresequenzen oder Nukleinsäurekonstrukte funktionell mit genetischen Kontrollsequenzen verbunden.

10 Beispiele erfindungsgemässer Vektoren können Expressionskonstrukte folgenden Typs umfassen:

a) 5'-Pflanzenspezifischer Promotor / anti-HGD / Terminator-3'

15 b) 5'-Pflanzenspezifischer Promotor / anti-MAAI/FAAH / Terminator-3'

c) 5'-Pflanzenspezifischer Promotor / pro-HG / Terminator-3'

20 d) 5'-Pflanzenspezifischer Promotor / pro-VitaminE / Terminator-3'

Ausdrücklich betrifft die Erfindung auch Vektoren, die in der Lage sind Polypeptide mit einer HGD, MAAI, oder FAAH Aktivität  
25 zu exprimieren. Die für diese Gene kodierenden Sequenzen stammen bevorzugt aus Pflanzen, Cyanobakterien, Moosen, Pilzen oder Algen. Besonders bevorzugt sind die durch für Polypeptide gemäss SEQ ID NO 3, 5 und 18 kodierenden Sequenzen.

30 Hierbei können die kodierende pro-HG oder pro-VitaminE Sequenz, sowie die zur Expression von Polypeptiden mit HGD, MAAI, oder FAAH Aktivität dienenden Sequenzen auch durch eine kodierende Sequenz für ein Fusionsprotein aus Transitpeptid und der entsprechenden Sequenz ersetzt sein.

35

Bevorzugte Beispiele umfassen Vektoren und können eines der folgenden Expressionskonstrukte enthalten:

a) 5'-35S-Promotor / anti-MAAI/FAAH / OCS-Terminator-3'

40

b) 5'-35S-Promotor / anti-HGD / OCS-Terminator-3';

c) 5'-LeguminB-Promotor / pro-HG / NOS-Terminator-3'

45 d) 5'-LeguminB-Promotor / pro-VitaminE / NOS-Terminator-3'

- e) 5'-LeguminB-Promotor / HGD / NOS-Terminator-3'
- f) 5'-LeguminB-Promotor / MAAI / NOS-Terminator-3'
- 5 g) 5'-LeguminB-Promotor / FAAH / NOS-Terminator-3'

Auch hierbei kann die kodierende pro-HG Sequenz oder pro-Vitamin E Sequenz auch durch eine kodierende Sequenz für ein Fusionsprotein aus Transitpeptid und pro-HG oder pro-VitaminE ersetzt  
10 sein.

Für die erfindungsgemässen vorteilhaften Verfahren zur Optimierung der Vitamin E Biosynthese kann eine Kotransformation mit mehr als einem der oben genannten Beispiele a.) bis g.)  
15 erforderlich sein. Ferner kann die Transformation mit einem oder mehr Vektoren, die jeweils eine Kombination der oben genannten Konstrukte enthalten, vorteilhaft sein. Bevorzugte Beispiele umfassen Vektoren, enthaltend folgende Konstrukte:

- 20 a) 5'-35S-Promotor/ anti-MAAI/FAAH / OCS-Terminator / LeguminB-Promotor / pro-HG / NOS-Terminator-3';
- b) 5'-35S-Promotor/ anti-MAAI/FAAH / OCS-Terminator / LeguminB-Promotor / pro-VitaminE / NOS-Terminator-3';
- 25 c) 5'-35S-Promotor/ anti-HGD / OCS-Terminator / LeguminB-Promotor / pro-VitaminE / NOS-Terminator-3';
- d) 5'-35S-Promotor/ pro-HG / OCS-Terminator / LeguminB-Promotor / pro-VitaminE / NOS-Terminator-3';
- 30

Die Konstrukte a) bis d) erlaubt die gleichzeitige Transformation der Pflanze mit pro-HG bzw. pro-VitaminE und anti-HGD bzw. anti-MAAI/FAAH .

35 Unter Verwendung der oben zitierten Rekombinations- und Klonierungstechniken können die Nukleinsäurekonstrukte in geeignete Vektoren kloniert werden, die ihre Vermehrung, beispielsweise in *E. coli*, ermöglichen. Geeignete Klonierungsvektoren sind u.a.  
40 pBR332, pUC-Serien, M13mp-Serien und pACYC184. Besonders geeignet sind binäre Vektoren, die sowohl in *E. coli* als auch in Agrobakterien replizieren können.

Die erfindungsgemässen Nukleinsäurekonstrukte werden bevorzugt in  
45 geeignete Transformationsvektoren insertiert. Geeignete Vektoren sind unter anderem in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Kap. 6/7, S. 71-119 (1993) beschrie-

ben. Vorzugsweise werden sie in einen Vektor, wie beispielsweise pBin19, pBinAR, pPZP200 oder pPTV, kloniert, der geeignet ist, Agrobacterium tumefaciens zu transformieren. Die mit einem solchen Vektor transformierten Agrobakterien können dann in 5 bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen, wie z.B. von Raps, verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschliessend in geeigneten Medien kultiviert werden. Die Transformation von Pflanzen durch Agro- 10 bakterien ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15 - 38. Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter 15 Weise transgene Pflanzen regeneriert werden, die integriert die oben beschriebenen Nukleinsäurekonstrukte enthalten.

Die in den erfindungsgemässen Nukleinsäurekonstrukten und Vektoren enthaltenen Nukleinsäuresequenzen können mit mindestens einer 20 genetischen Kontrollsequenz funktionell verknüpft sein. Genetische Kontrollsequenzen gewährleisten zum Beispiel die Transkription und Translation in prokaryontischen oder eukaryontischen Organismen. Vorzugsweise umfassen die erfindungsgemässen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz 25 einen Promotor und 3'-stromabwärts eine Terminatorsequenz, sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils funktionell verknüpft mit der kodierenden Sequenz. Unter einer funktionellen Verknüpfung versteht man zum Beispiel die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulativer Elemente derart, dass jedes 30 der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz oder der antisense-Sequenz bestimmungsgemäss erfüllen kann. Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, 35 wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von anderen DNA-Molekülen aus auf die Zielsequenz ausüben.

Beispiele sind Sequenzen, an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich 40 zu diesen neuen Kontrollsequenzen oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Das 45 Nukleinsäurekonstrukt kann aber auch einfacher aufgebaut sein, das heisst, es werden keine zusätzlichen Regulationssignale vor die vorstehend erwähnten Gene inseriert und der natürliche

Promotor mit seiner Regulation wird nicht entfernt. Stattdessen wird die natürliche Kontrollsequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert wird. Diese veränderten Promotoren können auch allein vor die natürlichen  
5 Gene zur Steigerung der Aktivität gebracht werden.

Das Nukleinsäurekonstrukt kann ausserdem vorteilhafterweise eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression  
10 der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen insertiert werden, wie weitere regulatorische Elemente oder Terminatoren. Die vorstehend erwähnten Gene können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

15 Zusätzliche zur funktionellen Verknüpfung bevorzugte aber nicht darauf beschränkte Sequenzen sind weitere, von den Transitpeptid kodierenden Sequenzen verschiedene, Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der  
20 Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten; sowie Translationsverstärker wie die 5'-Leadersequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693 -8711), und dergleichen.

25 Als Kontrollsequenzen sind weiterhin solche zu verstehen, die eine homologe oder heterologe Rekombination bzw. Insertion in das Genom eines Wirtsorganismus ermöglichen oder die Entfernung aus dem Genom erlauben. Bei der homologen Rekombination kann zum Beispiel das endogene Gen gänzlich inaktiviert werden. Es kann ferner durch ein synthetisches Gen mit erhöhter und veränderter Aktivität ausgetauscht werden. Methoden wie die cre/lox-Technologie erlauben eine gewebsspezifische, unter Umständen induzierbare Entfernung des Zielgens aus dem Genom des Wirtsorganismus (Sauer  
30 B. Methods. 1998; 14(4):381-92). Hier werden bestimmte flankierende Sequenzen dem Zielgen angefügt (lox-Sequenzen), die später eine Entfernung mittels der cre-Rekombinase ermöglichen.

Je nach nachstehend näher beschriebenen Wirtsorganismus oder Ausgangsorganismus, der durch Einbringen der Nukleinsäurekonstrukte  
40 in einen genetisch veränderten oder transgenen Organismus überführt wird, eignen sich verschiedene Kontrollsequenzen.

Vorteilhafte Kontrollsequenzen für die erfindungsgemässen  
45 Nukleinsäurekonstrukte, für die erfindungsgemässen Vektoren, für das erfindungsgemäße Verfahren zur Herstellung von Vitamin E und für die nachstehend beschriebenen genetisch veränderten Organis-

men sind beispielsweise in Promotoren wie *cos-*, *tac-*, *trp-*, *tet-*, *lpp-*, *lac-*, *lpp-lac-*, *lacIq-*, *T7-*, *T5-*, *T3-*, *gal-*, *trc-*, *ara-*, *SP6-*, *l-PR-* oder im *l-PL-Promotor* enthalten, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden.

5

Weitere vorteilhafte Kontrollsequenzen sind beispielsweise in den gram-positiven Promotoren *amy* und *SPO2*, in den Hefe- oder Pilz-promotoren *ADC1*, *MFa*, *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH* oder in den Pflanzenpromotoren *CaMV/35S* [Franck et al., Cell 21(1980)

10 285-294], *PRP1* [Ward et al., Plant. Mol. Biol. 22 (1993)], *SSU*, *OCS*, *LEB4*, *USP*, *STLS1*, *B33*, *NOS*; *FBPaseP* (WO 98/18940) oder im *Ubiquitin-* oder *Phaseolin-Promotor* enthalten.

Als bevorzugte Promotoren für die Nukleinsäurekonstrukte ist

15 grundsätzlich jeder Promotor geeignet, der die Expression von Genen, insbesondere Fremdgenen, in Pflanzen steuern kann.

Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Insbesondere bevorzugt ist der *CaMV 35S-Promotor* aus dem Blumen-

20 kohl-Mosaik-Virus [Franck et al., Cell 21 (1980), 285 - 294].

Dieser Promotor enthält bekanntlich unterschiedliche Erkennungssequenzen für transkriptionale Effektoren, die in ihrer Gesamtheit zu einer permanenten und konstitutiven Expression des eingeführten Gens führen [Benfey et al., EMBO J. 8 (1989), 2195-2202].

25 Ein weiteres Beispiel eines geeigneten Promotors ist der *LeguminB-Promotor* (Accessionnr. X03677).

Die Nukleinsäurekonstrukte können auch einen chemisch induzierbaren Promotor enthalten, durch den die Expression des exogenen

30 Gens in der Pflanze zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige Promotoren, wie z.B. der *PRP1 Promotor* [Ward et al., Plant. Mol. Biol. 22 (1993), 361-366], ein durch *Salicylsäure* induzierbarer (WO 95/19443), ein durch *Benzolsulfonamid*-induzierbarer (EP-A-0388186), ein durch *Tetrazyklin*-induzierbarer (Gatz et al., (1992) Plant J. 2, 397404), ein durch *Abscisinsäure*-induzierbarer (EP-A 335528) bzw. ein durch *Ethanol*- oder *Cyclohexanon*-induzierbarer (WO 93/21334) Promotor können ebenfalls verwendet werden.

40 Weiterhin sind insbesondere solche Promotoren bevorzugt, die die Expression in Geweben oder Pflanzenteilen sicherstellen, in denen die Biosynthese von Vitamin E bzw. dessen Vorstufen stattfindet oder in denen die Produkte vorteilhafterweise akkumuliert werden. Insbesondere zu nennen sind Promotoren für die ganze Pflanze auf-

45 grund konstitutiver Expression, wie beispielsweise der *CaMV Promotor*, der *OCS Promotor* aus *Agrobacterium* (*Octopin Synthase*), der *NOS Promotor* aus *Agrobacterium* (*Nopalin synthase*), der *Ubi-*

- quitin Promotor, Promotoren vakuolärer ATPase Untereinheiten oder der Promotor eines Prolin-reichen Proteins aus Weizen (WO 91/13991). Insbesondere zu nennen sind Promotoren, die eine blattspezifische Expression gewährleisten. Zu nennen sind der
- 5 Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubisco (Ribulose-1,5-bisphosphatcarboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al., EMBO J. 8 (1989), 2445 - 245). Beispiele für samenspezifische Promotoren sind der Phaseolin-Promotor
- 10 (US 5504200), der USP-Promotor (Baumlein, H. et al., Mol. Gen. Genet. (1991) 225 (3), 459 - 467) oder der LEB4-Promotor (Fiedler, U. et al., Biotechnology (NY) (1995), 13 (10) 1090) zusammen mit dem LEB4-Signalpeptid.
- 15 Weitere geeignete Promotoren sind beispielsweise spezifische Promotoren für Knollen, Speicherwurzeln oder Wurzeln, wie beispielsweise der Patatin Promotor Klasse I (B33), der Promotor des Cathepsin D Inhibitors aus Kartoffel, der Promotor der Stärke Synthese (GBSS1) oder der Sporamin Promotor, fruchtspezifische
- 20 Promotoren, wie beispielsweise der fruchtspezifische Promotor aus Tomate (EP-A 409625), fruchtreifung-spezifische Promotoren, wie beispielsweise der fruchtreifung-spezifische Promotor aus Tomate (WO 94/21794), blütenspezifische Promotoren, wie beispielsweise der Phytoen Synthese Promotor (WO 92/16635) oder der Promotor des
- 25 P-rr Gens (WO 98/22593) oder spezifische Plastiden- oder Chromoplasten-Promotoren, wie beispielsweise der RNA-Polymerase Promotor (WO 97/06250) oder auch der Promotor der Phosphoribosylpyrophosphat Amidotransferase aus Glycine max (siehe auch Genbank Accession Nummer U87999) oder ein anderer Nodien-spezifischer
- 30 Promotor wie in EP-A 249676 können vorteilhaft verwendet werden.

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für das erfindungsgemäße Verfahren verwendet werden. Darüberhinaus können auch syntheti-

35 sche Promotoren vorteilhaft verwendet werden.

- Als Kontrollsequenzen geeignete Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus Agrobacterium
- 40 tumefaciens, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACHS entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente davon. Beispiele für besonders geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopalinsynthase)-Ter-
- 45 minator.

Die Herstellung eines Nukleinsäurekonstrukts erfolgt beispielsweise durch Fusion eines geeigneten Promotors mit einer geeigneten anti-HGD, anti-MAAI/FAAH, pro-HG, pro-VitaminE, HGD, MAAI, oder FAAH Nukleotidsequenz, gegebenenfalls einer für eine Transitpeptid kodierenden Sequenz, vorzugsweise ein chloroplasten-spezifisches Transitpeptid, welche vorzugsweise zwischen dem Promotor und der jeweiligen Nukleotidsequenz angeordnet ist, sowie einem Terminator- oder Polyadenylierungssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.

Wie bereits erwähnt, können auch Nukleinsäurekonstrukte verwendet werden, deren DNA-Sequenz für ein pro-HG, pro-VitaminE, HGD, MAAI, oder FAAH Fusionsproteine kodiert, wobei ein Teil des Fusionsproteins ein Transitpeptid ist, das die Translokation des Polypeptides steuert. Als Beispiel können genannt werden: Chloroplasten-spezifische Transitpeptide, welche nach Translokation in die Chloroplasten enzymatisch abgespalten werden.

Bevorzugt werden die pro-HG, pro-VitaminE, HGD, MAAI, oder FAAH-Nukleotidsequenz mit der kodierenden Sequenz eines Pflanzenorganell-spezifischen Transitpeptids funktional verknüpft. Das Transitpeptid besitzt dabei vorzugsweise Spezifität für einzelne Zellkompartimente der Pflanze, zum Beispiel den Plastiden, wie zum Beispiel die Chloroplasten, Chromoplasten und/oder Leukoplasten. Das Transitpeptid lenkt die exprimierten Polypeptide an den gewünschten Zielort in der Pflanze und wird nach dessen Erreichen vorzugsweise proteolytisch abgespalten. Die kodierende Transitpeptid-Sequenz befindet sich im erfindungsgemässen Expressionskonstrukt vorzugsweise 5'-stromaufwärts von der kodierenden pro-HG, pro-VitaminE, HGD, MAAI, oder FAAH -Sequenz. Insbesondere ist zu nennen das Transitpeptid, das von der plastidären Nicotiana tabacum Transketolase (TK) oder einem funktionellen Äquivalent dieses Transitpeptids (z.B. dem Transitpeptid der kleinen Untereinheit der RubisCO oder der Ferredoxin:NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2) abgeleitet ist.

Ein anderer Gegenstand der Erfindung betrifft transgene Organismen, transformiert mit wenigstens einem erfindungsgemässen Nukleinsäurekonstrukt oder einem erfindungsgemässen Vektor, sowie



Zellen, Zellkulturen, Gewebe, Teile - wie zum Beispiel bei pflanzlichen Organismen Blätter, Wurzeln usw.- oder Vermehrungsgut abgeleitet von solchen Organismen.

- 5 Unter Organismus, Ausgangs- oder Wirtsorganismen werden prokaryontische oder eukaryontische Organismen, wie beispielsweise Mikroorganismen oder pflanzliche Organismen verstanden. Bevorzugte Mikroorganismen sind Bakterien, Hefen, Algen oder Pilze.
- 10 Bevorzugte Bakterien sind Bakterien der Gattung Escherichia, Erwinia, Agrobacterium, Flavobacterium, Alcaligenes oder Cyanobakterien zum Beispiel der Gattung Synechocystis.  
  
Bevorzug sind vor allem Mikroorganismen, welche zur Infektion  
15 von Pflanzen und damit zur Übertragung der erfindungsgemässen Konstrukte befähigt sind. Bevorzugte Mikroorganismen sind solche aus der Gattung Agrobacterium und insbesondere der Art Agrobacterium tumefaciens.
- 20 Bevorzugte Hefen sind Candida, Saccharomyces, Hansenula oder Pichia.  
  
Pflanzliche Organismen im Sinne der Erfindung sind mono- und dikotyle Pflanzen. Die erfindungsgemässen transgenen Pflanzen  
25 sind insbesondere ausgewählt unter monokotylen Kulturpflanzen, wie zum Beispiel Getreidearten wie Weizen, Gerste, Hirse, Roggen, Triticale, Mais, Reis oder Hafer, sowie dem Zuckerrohr. Ferner sind die erfindungsgemässen transgenen Pflanzen insbesondere ausgewählt unter dikotylen Kulturpflanzen, wie zum Beispiel  
30 Brassicaceae wie Raps, Kresse, Arabidopsis, Kohlarten oder Canola, Leguminosae wie Soja, Alfalfa, Erbse, Bohnengewächsen oder Erdnuss
- 35 Solanaceae wie Kartoffel, Tabak, Tomate, Aubergine oder Paprika, Asteraceae wie Sonnenblume, Tagetes, Salat oder Calendula, Cucurbitaceae wie Melone, Kürbis oder Zucchini, sowie Lein, Baumwolle, Hanf, Flachs, Roter Pfeffer, Karotte, Zuckerrübe und den verschiedenen Baum-, Nuss- und Weinspecies.  
40 Besonders bevorzug sind *Arabidopsis thaliana*, *Nicotiana tabacum*, *Tagetes erecta*, *Calendula vulgaris* sowie alle Gattungen und Arten, die sich zur Herstellung von Ölen eignen, wie Ölsaaten (wie zum Beispiel Raps), Nussarten, Soja, Sonnenblume, Kürbis  
45 und Erdnuss.

Pflanzliche Organismen im Sinne der Erfindung sind weiterhin weitere photosynthetisch aktive oder zur Vitamin E Synthese befähigte Organismen, wie zum Beispiel Algen oder Cyanobakterien, sowie Moose.

5

Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* oder *Dunaliella*.

- 10 Als Transformation wird die Übertragung von Fremdgenen in das Genom eines Organismus zum Beispiel einer Pflanze bezeichnet. Es werden dabei die beschriebenen Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt. Geeignete
- 15 Methoden sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone, die sogenannte particle bombardment Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikroinjektion und der durch *Agrobacterium* ver-
- 20 mittelte Gentransfer. Die genannten Verfahren sind beispielsweise in B. Jenes et al., *Techniques for Gene Transfer*, in: *Transgenic Plants*, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press (1993), 128 - 143 sowie in Potrykus, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991),
- 25 205 - 225) beschrieben. Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, *Agrobacterium tumefaciens* zu transformieren, beispielsweise pBin19 (Bevan et al., *Nucl. Acids Res.* 12 (1984), 8711).
- 30 Die Wirksamkeit der Expression der transgen exprimierten Nukleinsäuren kann beispielsweise *in vitro* durch Sprossmeristemvermehrung ermittelt werden. Zudem kann eine in Art und Höhe veränderte Expression von pro-HG oder pro-VitaminE Genen und deren Auswirkung auf die Vitamin E - Biosyntheseleistung an Testpflanzen in
- 35 Gewächshausversuchen getestet werden.

Ein weiterer Gegenstand der Erfindung sind transgene Organismen nach obiger Beschreibung, die zu einer im Vergleich zum untransformierten Wildtyp verbesserten Vitamin E Produktion befähigt

40 sind.

Erfindungsgemäss sind ferner von den oben beschriebenen transgenen Organismen abgeleitete Zellen, Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln,

45 Blätter etc. - , transgenes Vermehrungsgut, Saaten oder Früchte.

Eine verbesserte Vitamin E-Produktion bedeutet im Rahmen der vorliegenden Erfindung zum Beispiel die künstlich erworbene Fähigkeit einer erhöhten Biosyntheseleistung wenigstens einer Verbindung aus der Gruppe der Tocopherole und Tocotrienole, 5 in dem transgenen Organismen gegenüber dem nicht gentechnisch modifizierten Ausgangsorganismus für die Dauer mindestens einer Pflanzengeneration. Dabei ist die Vitamin E-Produktion in dem transgenen Organismus gegenüber dem nicht-gentechnisch modifizierten Ausgangsorganismus bevorzugt um 10 %, besonders bevorzugt 10 um 50 %, ganz besonders bevorzugt um 100 % erhöht. Verbessert kann ebenfalls eine vorteilhaft veränderte qualitative Zusammensetzung des VitaminE-Gemisches bedeuten.

Der Biosyntheseort von Vitamin E ist im allgemeinen das Blatt- 15 gewebe aber auch der Samen, so dass eine blattspezifische oder samenspezifische Expression insbesondere von pro-HG und pro-Vitamin E Sequenzen und gegebenenfalls von anti-HGD bzw. anti-MAAI/FAAH Sequenzen sinnvoll ist. Es ist jedoch naheliegend, dass die Vitamin E-Biosynthese nicht auf den Samen beschränkt sein muss, 20 sondern auch in allen übrigen Teilen der Pflanze gewebespezifisch erfolgen kann. Darüberhinaus ist eine konstitutive Expression des exogenen Gens von Vorteil. Andererseits kann aber auch eine induzierbare Expression wünschenswert sein.

25 Ein weiterer Gegenstand der Erfindung betrifft schliesslich ein Verfahren zur Herstellung von Vitamin E, das dadurch gekennzeichnet ist, dass man aus einer Kultur eines erfindungsgemäss transformierten pflanzlichen Organismus das gewünschte Vitamin E in an sich bekannter Weise isoliert.

30 Von Menschen und Tieren verzehrbare erfindungsgemässe, genetisch veränderte Pflanzen mit erhöhtem VitaminE-Gehalt können auch beispielsweise direkt oder nach an sich bekannter Aufbereitung als Nahrungsmittel oder Futtermittel verwendet werden.

35 Die Erfindung betrifft ferner die Verwendung von Polypeptiden, die für eine HGD, MAAI oder FAAH kodieren, der ihnen zugrunde liegenden Gene und cDNAs, bzw. der von ihnen abgeleiteten erfindungsgemässen Nukleinsäurekonstrukte, erfindungsgemässen Vektoren 40 oder erfindungsgemässen Organismen zur Herstellung von Antikörpern, protein- oder DNA-bindenden Faktoren.

Der Biosyntheseweg von HGD-MAAI-FAAH-Abbauweg bietet Targetenzyme für die Entwicklung von Inhibitoren. Daher betrifft die Erfindung 45 auch die Verwendung von Polypeptiden, die für eine HGD, MAAI oder FAAH kodieren, der ihnen zugrunde liegenden Gene und cDNAs, bzw. der von ihnen abgeleiteten erfindungsgemässen Nukleinsäurekon-

strukture, erfindungsgemässen Vektoren oder erfindungsgemässen Organismen als Target zum Auffinden von Inhibitoren der HGD, MAAI oder FAAH.

- 5 Um effiziente Hemmstoffe der HGD, MAAI oder FAAH finden zu können, ist es notwendig, geeignete Testsysteme, mit denen Inhibitor-Enzym-Bindungsstudien durchgeführt werden können, zur Verfügung zu stellen. Hierzu wird beispielsweise die komplette cDNA-Sequenz der HGD, MAAI oder FAAH in einen Expressionsvektor (zum Beispiel pQE, Qiagen) kloniert und in *E. coli* überexprimiert. Die HGD, MAAI oder FAAH -Proteine eignet sich besonders zur Auffindung von für die HGD, MAAI oder FAAH spezifischen Hemmstoffen.

- Dementsprechend betrifft die Erfindung ein Verfahren zum Auffinden von Inhibitoren der HGD, MAAI oder FAAH unter Verwendung von oben genannten Polypeptiden, Nukleinsäuren, Vektoren oder transgenen Organismen, dadurch gekennzeichnet, dass man die enzymatische Aktivität der HGD, MAAI oder FAAH in Gegenwart einer chemischen Verbindung misst und bei Erniedrigung der enzymatischen Aktivität im Vergleich zur nicht gehemmten Aktivität die chemische Verbindung einen Inhibitor darstellt. Dazu kann die HGD, MAAI oder FAAH beispielsweise in einem Enzymtest eingesetzt werden, bei dem die Aktivität der HGD, MAAI oder FAAH in An- und Abwesenheit des zu testenden Wirkstoffs ermittelt wird. Aus dem Vergleich der beiden Aktivitätsbestimmungen lässt sich eine qualitative und quantitative Aussage über das Hemmverhalten des zu testenden Wirkstoffes machen. Mit Hilfe des erfindungsgemässen Testsystems kann eine Vielzahl von chemischen Verbindungen schnell und einfach auf herbizide Eigenschaften überprüft werden.
- 30 Das Verfahren gestattet es, reproduzierbar aus einer grossen Anzahl von Substanzen gezielt solche mit grosser Wirkstärke auszuwählen, um mit diesen Substanzen anschliessend weitere, dem Fachmann geläufige vertiefte Prüfungen durchzuführen.

- 35 Die Inhibitoren der HGD, MAAI oder FAAH eignen sie sich zur Steigerung der Vitamin E-Biosynthese funktionell analog den oben beschriebenen anti-HGD bzw. anti-MAAI/FAAH Nukleinsäuresequenzen. Ein weiterer Gegenstand der Erfindung sind daher Verfahren zur Verbesserung der Vitamin E Produktion unter Verwendung von Inhibitoren der HGD, MAAI oder FAAH. Die verbesserte Produktion von Vitamin E kann einen positiven Effekt auf die Pflanze bewirken, da diese Verbindungen eine wichtige Funktion im Schutz vor schädlichen Umwelteinflüssen (Sonnenstrahlung, Sauerstoffradikale) haben. Eine Steigerung der Vitamin E Produktion kann insofern als Wachstumsförderer fungieren. Ein weiterer Gegenstand der Erfindung ist daher die Verwendung von Inhibitoren der HGD, MAAI oder

FAAH, erhältlich durch dem oben beschriebenen Verfahren, als Wachstumsregulatoren.

## Sequenzen

5

- SEQ ID NO. 1: Homogentisat-1,2-dioxygenase (HGD) Gen aus *Arabidopsis thaliana*
- SEQ ID NO. 2: Homogentisat-1,2-dioxygenase (HGD) cDNA aus *Arabidopsis thaliana*
- 10 SEQ ID NO. 3: Homogentisat-1,2-dioxygenase (HGD) Polypeptid aus *Arabidopsis thaliana*
- SEQ ID NO. 4: Fumarylacetoacetathydrolase (FAAH) cDNA aus *Arabidopsis thaliana*
- SEQ ID NO. 5: Fumarylacetoacetathydrolase (FAAH) Polypeptid aus *Arabidopsis thaliana*
- 15 SEQ ID NO. 6: Maleylacetoacetatisomerase (MAAI) Gen aus *Arabidopsis thaliana*
- SEQ ID NO. 7: TyrA Gen kodierend für eine bifunktionale Chorismatmutase/ Prephenatdehydrogenase
- 20 SEQ ID NO. 8: TyrA Polypeptid kodierend für eine bifunktionale Chorismatmutase/ Prephenatdehydrogenase
- SEQ ID NO. 9: Fumarylacetoacetathydrolase (FAAH) Gen aus *Arabidopsis thaliana*
- SEQ ID NO. 10: Hydroxyphenylpyruvatdioxygenase (HPPD) cDNA aus *Arabidopsis thaliana*
- 25 SEQ ID NO. 11: Hydroxyphenylpyruvatdioxygenase (HPPD) Polypeptid aus *Arabidopsis thaliana*
- SEQ ID NO. 12: Homogentisat-1,2-dioxygenase (HGD) cDNA Fragment aus *Brassica napus*
- 30 SEQ ID NO. 13: Homogentisatphythyltransferase cDNA aus *Synechocystis* PCC6803
- SEQ ID NO. 14: Homogentisatphythyltransferase Polypeptid aus *Synechocystis* PCC6803
- SEQ ID NO. 15: Künstliche codonusage optimierte cDNA kodierend für Hydroxyphenylpyruvatdioxygenase (HPPDop) aus *Streptomyces avermitilis*
- 35 SEQ ID NO. 16: Hydroxyphenylpyruvatdioxygenase Polypeptid aus *Streptomyces avermitilis*
- SEQ ID NO. 17: Maleylacetoacetatisomerase (MAAI) cDNA aus *Arabidopsis thaliana*
- 40 SEQ ID NO. 18: Maleylacetoacetatisomerase (MAAI) Polypeptid aus *Arabidopsis thaliana*
- SEQ ID NO. 19:  $\gamma$ -Tocopherolmethyltransferase cDNA aus *Arabidopsis thaliana*
- 45 SEQ ID NO. 20:  $\gamma$ -Tocopherolmethyltransferase Polypeptid aus *Arabidopsis thaliana*

- SEQ ID NO. 21: 3-Methyl-6-phytylhydrochinonmethyltransferase cDNA aus *Synechocystis* PCC6803
- SEQ ID NO. 22: 3-Methyl-6-phytylhydrochinonmethyltransferase Polypeptid aus *Synechocystis* PCC6803
- 5 SEQ ID NO. 23: Geranylgeranylpyrophosphatooxidoreduktase cDNA aus *Nicotiana tabacum*.
- SEQ ID NO. 24: Geranylgeranylpyrophosphatooxidoreduktase Polypeptid aus *Nicotiana tabacum*.
- SEQ ID NO. 25: Primer (5'-HGD *Brassica napus*)  
10 5'-GTCGACGGNCCNATNGGNGCNAANGG-3'
- SEQ ID NO. 26: Primer (3'-NOS Terminator)  
5'-AAGCTTCCGATCTAGTAACATAGA-3'
- SEQ ID NO. 27: Primer (5'-35 S Promotor)  
5'-ATTCTAGACATGGAGTCAAAGATTCAAATAGA-3'
- 15 SEQ ID NO. 28: Primer (3'-OCS Terminator)  
5'-ATTCTAGAGGACAATCAGTAAATTGAACGGAG-3'
- SEQ ID NO. 29: Primer (5'-MAAI *A.thaliana*)  
5'-atgtcgacATGTCTTATGTTACCGAT-3'
- SEQ ID NO. 30: Primer (3'-MAAI *A.thaliana*)  
20 5'-atggatccCTGGTTCATATGATACA-3'
- SEQ ID NO. 31: Primer (5'-FAAH *A.thaliana*)  
5'-atgtcgacGGAACTCTGAACCATAT-3'
- SEQ ID NO. 32: Primer (3'-FAAH *A.thaliana*)  
5'-atggtaccGAATGTGATGCCTAAGT-3'
- 25 SEQ ID NO. 33: Primer (3'-HGD *Brassica napus*)  
5'-GGTACCTCRAACATRAANGCCATNGTNCC-3'
- SEQ ID NO. 34: Primer (5'-Legumin Promotor)  
5'-GAATTCGATCTGTCGTCTCAAACCTC-3'
- SEQ ID NO. 35: Primer (3'-Legumin Promotor)  
30 5'-GGTACCGTGATAGTAAACAATAATG-3'
- SEQ ID NO. 36: Primer (5'-Transitpeptid)  
5'-ATGGTACCTTTTTTGCATAAACTTATCTTCATAG-3'
- SEQ ID NO. 37: Primer (3'-Transitpeptid)  
5'-ATGTCGACCCGGGATCCAGGGCCCTGATGGGTCCCATTTTCCC-3'
- 35 SEQ ID NO. 38: Primer (5'-NOS Terminator)  
5'-GTCGACGAATTTCCCGAATCGTTC-3'
- SEQ ID NO. 39: Primer (3'-NOS Terminator II)  
5'-AAGCTTCCGATCTAGTAACATAGA-3'
- SEQ ID NO. 40: Primer (5'-Legumin Promotor II)  
40 5'-AAGCTTGATCTGTCGTCTCAAACCTC-3'
- SEQ ID NO. 41: Maleylacetoacetatisomerase (MAAI) Gen (Fragment) aus *Arabidopsis thaliana*
- SEQ ID NO. 42: Fumarylacetoacetathydrolase (FAAH) Gen (Fragment) aus *Arabidopsis thaliana*
- 45 SEQ ID NO. 43: Primer (5'-35 S Promotor)  
5'-ATGAATTCATGGAGTCAAAGATTCAAATAGA-3'

SEQ ID NO. 44: Primer (3'-OCS Terminator)

5'-ATGAATTCGGACAATCAGTAAATTGAACGGAG-3'

### Beispiele

5

Die Erfindung wird in den folgenden Ausführungsbeispielen unter Bezugnahme auf die beiliegenden Figuren näher erläutert. Dabei werden Abkürzungen mit folgender Bedeutung verwendet:

- |                                       |                                    |
|---------------------------------------|------------------------------------|
| 10 A = 35S-Promotor                   | B = HGD in antisense-Orientierung  |
| C = OCS Terminator                    | D = Legumin B-Promotor             |
| E = Transitpeptid der FNR             | F = HPPDop                         |
|                                       | (HPPD mit optimierter Codonusage)  |
| G = NOS-Terminator                    | H = MAAI in antisense-Orientierung |
| 15 I = FAAH in antisense-Orientierung |                                    |

Die Richtung von Pfeilen in den Figuren zeigt jeweils der Verlauf der Leserichtung der entsprechenden Gene an. Dabei zeigt:

- 20 Figur 1 eine schematische Darstellung des Vitamin E - Biosyntheseweges in Pflanzen;
- Figur 2 Konstruktionsschemata der antiHGD kodierenden Plasmide pBinARHGDanti (I) und pCRScriptHGDanti (II);
- 25 Figur 3 Konstruktionsschemata der HPPDop kodierenden Plasmide pUC19HPPDop (III) und pCRScriptHPPDop (IV);
- Figur 4 Konstruktionsschemata der Transformationsvektoren pPTVHGDanti (V) und des bifunktionalen Transformationsvektors pPTV HPPDop HGD anti (VI), welcher die HPPDop in Samen transformierter Pflanzen exprimiert und gleichzeitig die Expression der endogenen HGD unterdrückt.
- 30
- 35 Figur 5 Konstruktionsschema des Transformationsvektors pZP200HPPDop (VII).
- Figur 6 Konstruktionsschemata der Transformationsvektoren pGEMT MAAI1 anti (VIII) und pBinAR MAAI1 anti (IX)
- 40
- Figur 7 Konstruktionsschemata der Transformationsvektoren pCR-Script MAAI1 anti (X) und pZPNBN MAAI1 anti (XI)
- Figur 8 Konstruktionsschemata der Transformationsvektoren pGEMT FAAH anti (XII)
- 45

Figur 9 Konstruktionsschemata der Transformationsvektoren pBinAR  
FAAH anti (XIII) und pZPNBN FAAH anti (XIV)

Allgemeine Methoden:

5 Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die im Rahmen der vorliegenden Erfindung durchgeführten Klonierungsschritte wie z.B. Restriktionsspaltungen, AgaroseGelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrozellulose und Nylonmembranen, Verknüpfen von DNA-Fragmenten, Transformation von *E. coli* Zellen, Anzucht von Bakterien, Vermehrung von Phagen und Sequenzanalyse rekombinanter DNA wurden wie bei Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6 beschrieben durchgeführt. Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma Licor (Vertrieb durch MWG Biotech, Ebersbach) nach der Methode von Sanger (Sanger et al., 20 Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Beispiel 1:

Klonierung einer Hydroxyphenylpyruvat-Dioxygenase (HPPD) mit für Expression in *Brassica napus* optimierter DNA-Sequenz

25 Die Aminosäuresequenz der Hydroxyphenylpyruvat-Dioxygenase (HPPD) aus *Streptomyces avermitilis* (Accessionnr. U11864, SEQ ID NO:16) wurde unter Berücksichtigung der Codonverwendung in *Brassica napus* (Raps) in eine DNA-Sequenz zurückübersetzt. Die Codonusage wurde mittels der Datenbank <http://www.dna.affrc.go.jp/~nakamura/index.html> bestimmt. Die abgeleitete Sequenz wurde unter Anheftung von SalI Schnittstellen durch Ligation überlappender Oligonukleotide mit anschliessender PCR-Amplifikation (Rouwendal, GJA; et al, (1997) PMB 33: 989-999) synthetisiert (SEQ ID NO:15). 35 Die Richtigkeit der Sequenz des synthetischen Gens wurde durch Sequenzierung überprüft. Das synthetische Gen wurde in den Vektor pBluescript II SK+ (Stratagene) kloniert. (Diese kodonoptimierte Sequenz ist infolge auch als HPPDop bezeichnet.)

40

45



## Beispiel 2:

Klonierung einer Homogentisat-Dioxygenase (HGD) aus *Brassica napus*

5 a) Isolierung von gesamt-RNA aus Blüten von *Brassica napus*

Von *Brassica napus* var. Westa wurden offene Blüten geerntet und in flüssigem Stickstoff eingefroren. Das Material wurde anschliessend im Mörser pulverisiert und in Z6-Puffer (8 M Guanidinium-Hydrochlorid, 20 mM MES, 20 mM EDTA, auf pH 7,0 mit NaOH eingestellt; versetzt mit 400 ml Mercaptoethanol/100 ml Puffer unmittelbar vor Gebrauch) aufgenommen. Die Suspension wurde dann in Reaktionsgefässe überführt und mit einem Volumen Phenol/Chloroform/Isoamylalkohol 25:24:1 ausgeschüttelt. Nach 10 minütiger Zentrifugation bei 15000 U wurde der Überstand in ein neues Reaktionsgefäss überführt und mit 1/20 Volumen 1N Essigsäure und 0,7 Volumen Ethanol (absolut) die RNA gefällt. Nach erneuter Zentrifugation wurde das Pellet zunächst in 3M Natriumacetatlösung und nach einer weiteren Zentrifugation in 70 % Ethanol gewaschen. Anschliessend wurde das Pellet in DEPC (Diethylpyrocarbonat) Wasser gelöst und die RNA-Konzentration photometrisch bestimmt.

b) Herstellung von cDNA aus gesamt RNA aus Blüten von *Brassica napus*

20 mg Gesamt-RNA wurden zunächst mit 3,3 ml 3M Natriumacetatlösung, 2 ml 1M Magnesiumsulfatlösung versetzt und auf 10 ml Endvolumen mit DEPC Wasser aufgefüllt. Dazu wurde 1 ml RNase-freie DNase (Boehringer Mannheim) gegeben und 45 min bei 37 Grad inkubiert. Nach Entfernen des Enzyms durch Ausschütteln mit Phenol/Chloroform/Isoamylalkohol wurde die RNA mit Ethanol gefällt und das Pellet in 100 ml DEPC Wasser aufgenommen. 2,5 mg RNA aus dieser Lösung wurden mittels eines cDNA-Kits (Gibco BRL) nach Herstellerangaben in cDNA umgeschrieben.

35

c) PCR-Amplifikation eines Teilfragments der HGD aus *Brassica napus*

Durch Vergleich der DNA-Sequenzen der bekannten Homogentisat-Dioxygenasen (HGD) aus *Arabidopsis thaliana* (Accessionnr. U80668), *Homo sapiens* (Accessionnr. U63008) und *Mus musculus* (Accessionnr. U58988) wurden für eine PCR Oligonukleotide abgeleitet, denen am 5'-Ende eine SalI und am 3'-Ende eine Asp718 Restriktionsschnittstelle angefügt worden war. Das Oligonukleotid am 5'-Ende umfasst die Sequenz:

5'-GTCGACGGNCCNATNGGNGCNAANGG-3' (SEQ ID NO:25),

beginnend mit der Base 661 des Arabidopsis-Gens. Das Oligonukleotid am 3'-Ende umfasst die Sequenz:

5

5'-GGTACCTCRAACATRAANGCCATNGTNCC-3' (SEQ ID NO:33),

beginnend mit der Base 1223 des Arabidopsis-Gens, wobei N jeweils Inosin bedeutet und R für den Einbau von A oder G in das Oligonukleotid steht.

10

Die PCR-Reaktion wurde mit der Taq-Polymerase von TAKARA nach Herstellerangaben durchgeführt. Als Template wurden 0,3 mg der cDNA eingesetzt. Das PCR-Programm lautete:

15

- 1Zyklus mit: 94°C (1 min)  
5Zyklen mit: 94°C (4 sec), 50°C (30 sec), 72°C (1 min)  
5Zyklen mit: 94°C (4 sec), 48°C (30 sec), 72°C (1 min)  
25Zyklen mit: 94°C (4 sec), 46 Grad (30 sec), 72 Grad (1 min)

20

- 1Zyklus mit: 72 Grad (30 min)

Das Fragment wurde mittels NucleoSpin Extract (Machery und Nagel) gereinigt und nach Herstellerangaben in den Vektor pGEMT (Promega) kloniert. Die Richtigkeit des Fragments wurde durch Sequenzierung überprüft.

25

Beispiel 3: Herstellung eines Pflanzentransformations-Konstrukts zur Überexpression der HPPD mit optimierter DNA-Sequenz (HPPDop) und Ausschaltung der HGD

30

Zur Herstellung von Pflanzen, welche die HPPDop in Samen exprimieren und in denen die Expression der endogenen HGD mittels antisense-Technik unterdrückt ist, wurde ein binärer Vektor angefertigt, der beide Gensequenzen enthält (Figur 4, Konstrukt VI).

35

a) Herstellung einer HPPDop- Nukleinsäurekonstrukt

Dazu wurden zunächst die Komponenten der Kassette zur Expression der HPPDop, bestehend aus dem LeguminB-Promotor (Accessionnr. X03677), dem Transitpeptid der Ferredoxin:NADP+ Oxidoreduktase aus Spinat (FNR; Jansen, T, et al (1988) Current Genetics 13, 517-522) und dem NOS-Terminator (enthalten im pBI101 Accessionnr. U12668) mittels PCR mit den benötigten Restriktionsschnittstellen versehen.

45

Der Legumin-Promotor wurde aus dem Plasmid plePOCS (Bäumlein, H, et al. (1986) Plant J. 24, 233-239) mit dem stromaufwärts-Oligonukleotid:

5 5'-GAATTCGATCTGTCGTCTCAAACCTC-3' (SEQ ID NO: 34)

und dem stromabwärts-Oligonukleotid:

5'-GGTACCGTGATAGTAAACAATAATG-3' (SEQ ID NO: 35)

10

mittels PCR amplifiziert und in den Vektor PCR-Script (Stratagene) nach Herstellerangaben kloniert.

Das Transitpeptid wurde aus dem Plasmid pSK-FNR (Andrea Babette  
15 Regierer "Molekulargenetische Ansätze zur Veränderung der  
Phosphat-Nutzungseffizienz von höheren Pflanzen", P+H Wissenschaftlicher Verlag, Berlin 1998 ISBN: 3-9805474-9-3) mittels PCR  
mit dem 5'-Oligonukleotid:

20 5'-ATGGTACCTTTTTTGCATAAACTTATCTTCATAG-3' (SEQ ID NO: 36)

und dem 3'-Oligonukleotid:

5'-ATGTCGACCCGGGATCCAGGGCCCTGATGGGTCCCATTTCCTCC-3' (SEQ ID NO: 37)

25

amplifiziert.

Der NOS-Terminator wurde aus dem Plasmid pBI101 (Jefferson, R.A., et al (1987) EMBO J. 6 (13), 3901-3907) mittels PCR mit dem  
30 5'-Oligonukleotid:

5'-GTCGACGAATTTCCCGAATCGTTC-3' (SEQ ID NO: 38)

und dem 3'-Oligonukleotid

35

5'-AAGCTTCCGATCTAGTAACATAGA-3' (SEQ ID NO: 26)

amplifiziert.

40 Das Amplikon wurde jeweils in den Vektor pCR-Script (Stratagene) nach Herstellerangaben kloniert.

Für die Nukleinsäurekonstruktion wurde zunächst der NOS-Terminator als SalI/HindIII-Fragment in einen entsprechend geschnittenen  
45 pUC19-Vektor (Yanisch-Perron, C., et al (1985) Gene 33, 103-119) umklont. In dieses Plasmid wurde anschliessend das Transitpeptid als Asp718/SalI-Fragment eingeführt. Der Legumin-Promotor

wurde dann als EcoRI/Asp718 Fragment einkloniert. Das Gen HPPDop wurde als Sali-Fragment in dieses Konstrukt eingeführt (Figur 3, Konstrukt III).

- 5 Die fertige Kassette in pUC19 wurde als Template für eine PCR verwendet, wozu für den Leguminpromotor das Oligonukleotid:

5'-AAGCTTGATCTGTCGTCTCAAAC-3' (SEQ ID NO: 40)

- 10 und für den Nos-Terminator das Oligonukleotid:

5'-AAGCTTCCGATCTAGTAACATAGA-3' (SEQ ID NO: 39)

- verwendet wurden. Das Amplikon wurde in pCR-Script kloniert und  
15 pCR-ScriptHPPDop genannt (Figur 3, Konstrukt IV).

d) Herstellung einer antiHGD-Nukleinsäurekonstrukt

- Für die Ausschaltung der HGD mit antisense-Technik wurde das Gen-  
20 fragment als Sali/Asp718-Fragment in den Vektor pBinAR (Höfgen, R. und Willmitzer, L., (1990) Plant Sci. 66: 221-230) kloniert, in dem der 35S-Promotor und der OCS-Terminator vorliegen (Figur 2, Konstrukt I). Das Konstrukt diente als Vorlage für eine PCR Reaktion mit dem Oligonukleotid:

25

5'-ATTCTAGACATGGAGTCAAAGATTCAAATAGA-3' (SEQ ID NO: 27),

spezifisch für die 35S-Promotor-Sequenz;

- 30 und dem Oligonukleotid:

5'-ATTCTAGAGGACAATCAGTAAATTGAACGGAG-3' (SEQ ID NO: 28).

spezifisch für OCS-Terminator-Sequenz

35

Das Amplikon wurde in den Vektor pCR-Script (Stratagene) kloniert und pCRScriptHGDanti genannt (Figur 2, Konstrukt II).

c) Herstellung des binären Vektors

40

- Zur Erstellung eines binären Vektors zur Raps-Transformation wurde zunächst das Konstrukt HGDanti aus pCRScriptHGDanti als XbaI-Fragment in den Vektor pPTV (Becker, D., (1992) PMB 20, 1195-1197) kloniert (Figur 4, Konstrukt V). In dieses Plasmid  
45 wurde das Konstrukt LegHPPDop aus pCRScriptHPPDop als HindIII-

Fragment eingefügt. Dieses Plasmid wurde mit pPTVHPPDopHGDanti bezeichnet (Figur 4, Konstrukt VI).

Beispiel 4:

- 5 Herstellung von Konstrukten zur Kotransformation zur Überexpression von HPPDop und Ausschaltung von HGD in *Brassica napus* Pflanzen

Zur Kotransformation von Pflanzen mit HPPDop und antiHGD wurde  
10 das Konstrukt LeguminB-Promotor/Tansitpeptid/HPPDop/NOS aus dem Vektor pCRScriptHPPDop (Figur 3, Konstrukt IV) als HindIII-Fragment herausgeschnitten und in den entsprechend geschnittenen Vektor pPZP200 (Hajdukiewicz, P., et al., (1994) PMB 25(6): 989-94) eingefügt (Figur 5, Konstrukt VII). Dieses Plasmid diente später  
15 zur Kotransformation von Pflanzen zusammen mit dem Vektor pPTVHGDanti (Figur 4, Konstrukt V) aus Beispiel 3 c).

Beispiel 5:

Klonierung eines genomischen Fragments der Maleylacetoacetat-Iso-  
20 merase aus *Arabidopsis thaliana*

a) Isolierung von genomischer DNA aus Blättern von *A. thaliana*:

Der verwendete Extraktionspuffer hat folgende Zusammensetzung:

25

- 1 Volumen DNA-Extraktionspuffer (0,35 M Sorbitol, 0,1 M Tris, 5 mM EDTA, pH 8,25 HCl)
- 1 Volumen Nuclei-Lysispuffer (0,2 M Tris-HCl pH 8,0, 50 mM  
30 EDTA, 2 M NaCl, 2% Hexadecyltrimethylammoniumbromide (CTAB))
- 0,4 Volumen 5% Natriumsarkosyl
- 0,38 g/100 ml Natriumbisulfit

35

100 mg Blattmaterial von *A. thaliana* wurden geerntet und in flüssigem Stickstoff eingefroren. Das Material wurde anschliessend im Mörser pulverisiert und in 750 µl Extraktionspuffer aufgenommen. Das Gemisch wurde 20 min bei 65°C erhitzt und anschliessend mit  
40 einem Volumen Chloroform/Isoamylalkohol (24:1) ausgeschüttelt. Nach 10 minütiger Zentrifugation bei 10000 rpm in einer Heraeus pico-fuge wurde der Überstand mit einem Volumen Isopropanol versetzt und die so gefällte DNA erneut 5 Minuten bei 10000 rpm pelletiert. Das Pellet wurde in 70 %igem Ethanol gewaschen, bei  
45 Raumtemperatur 10 min getrocknet und anschliessend in 100 µl TE-

RNase Puffer (10 mM Tris HCl pH 8.0, 1 mM EDTA pH 8.0, 100 mg/l RNase) gelöst.

b) Klonierung des Gens für die MAAI aus *Arabidopsis thaliana*

5

Mittels der Protein-Sequenz der MAAI aus Maus (*Mus musculus*) mittels BLAST-Suche in der NCBI Datenbank (<http://www.ncbi.nlm.nih.gov/BLAST/>) das MAAI-Gen aus *A. thaliana* identifiziert (Genbank Acc.-No. AAC78520.1). Die Sequenz ist in 10 der Genbank als putative Glutathione-S-Transferase annotiert.

Mittels der ID-Nummern der Proteinsequenz konnten die korrespondierende DNA-Sequenz ermittelt werden und Oligonukleotide abgeleitet werden. Den Oligonukleotiden wurde jeweils am 5'Ende eine SalI und am 3'Ende eine BamHI Restriktionsschnittstelle angefügt.

15 Das Oligonukleotid am 5'Ende umfasst die Sequenz

5'-atgtcgacATGTCCTTATGTTACCGAT-3' (SEQ ID NO: 29)

beginnend mit Base 37 der cDNA, dem ersten Codon, das Oligo-

20 nukleotid am 3'Ende umfasst die Sequenz

5'-atggatccCTGGTTCATATGATACA-3' (SEQ ID NO: 30)

beginnend mit dem Basenpaar 803 der cDNA-Sequenz. Die PCR-Reak-

25 tion wurde mit der Taq Polymerase (Hersteller: TaKaRa Shuzo Co., Ltd.) durchgeführt. Der Ansatz hatte folgende Zusammensetzung: 10 µl Puffer (20 mM Tris-HCl pH 8,0, 100 mM KCl, 0,1 mM EDTA, 1 mM DTT, 0,5 % Tween20, 0,5 % Nonidet P-40, 50 % Glycerol), jeweils 100 pmol der beiden Oligonukleotide, jeweils 20 nM an dATP, 30 dCTP, dGTP, dTTP, 2,5 Einheiten Taq-Polymerase, 1 µg genomische DNA, destilliertes Wasser add 100 µl. Das PCR-Programm lautete:

- 5 Zyklen mit: 94°C (4 sec), 52°C (30 sec), 72°C (1 min)
- 5 Zyklen mit: 94°C (4 sec), 50°C (30 sec), 72°C (1 min)
- 35 • 25 Zyklen mit: 94°C (4 sec), 48°C (30 sec), 72°C (1 min)

Das amplifizierte Fragment (SEQ ID NO: 41) wurde mittels Nucleo-Spin Extract (Machery-Nagel) gereinigt und nach Herstellerangaben in den Vektor pGEMTeasy von Promega kloniert (Figur 6, Konstrukt

40 VIII). Die Richtigkeit des Fragments wurde durch Sequenzierung überprüft. Mittels der durch die Primer an die Sequenz angefügten Restriktionsschnittstellen wurde das Gen als SalI/BamHI-Fragment in den entsprechend geschnittenen Vektor BinAR (Höfgen, R. und Willmitzer, L., (1990) Plant Sci. 66: 221-230) kloniert (Figur 6, 45 Konstrukt IX). Dieser enthält den 35S-Promotor des Blumenkohlmo-

saikvirus und die OCS-Terminationssequenz. Das Konstrukt diente als Vorlage für eine PCR Reaktion mit dem Oligonukleotid

5'-ATGAATTCCATGGAGTCAAAGATTCAAATAGA-3' (SEQ ID NO: 43)

5

spezifisch für die 35S-Promotor-Sequenz und dem Oligonukleotid

5'-ATGAATTCGGACAATCAGTAAATTGAACGGAG-3' (SEQ ID NO: 44)

- 10 spezifisch für den OCS-Terminator. Beiden Oligonukleotiden wurde eine EcoRI-Erkennungssequenz angefügt. Die PCR wurde mit der Pfu-Polymerase durchgeführt (Hersteller: Stratagene). Der Ansatz hatte folgende Zusammensetzung: 10 µl Puffer (200 mM Tris HCl pH 8.8, 20 mM MgSO<sub>4</sub>, 100 mM KCl, 100 mM Ammoniumsulfat, 1 % Triton X-100, 1 g/l Nuclease freies BSA), jeweils 100 pmol der beiden Oligonukleotide, jeweils 20 nM an dATP, dCTP, dGTP, dTTP, 2.5 Einheiten Pfu-Polymerase, 1 ng Plasmid DNA, destilliertes Wasser add 100 µl. Das PCR Programm lautete:

- 20 • 5 Zyklen mit: 94°C (4 sec), 52°C (30 sec), 72°C (2 min)
- 5 Zyklen mit: 94°C (4 sec), 50°C (30 sec), 72°C (2 min)
- 25 Zyklen mit: 94°C (4 sec), 48°C (30 sec), 72°C (2 min)

- Das PCR-Fragment wurde mittels Nucleo-Spin Extract (Machery-  
25 Nagel) gereinigt und in den Vektor pCR-Script (Stratagene) kloniert (Figur 7, Konstrukt X).

#### Beispiel 6: Herstellung des binären Vektors

- 30 Zur Erstellung eines binären Vektors zur Arabidopsis- und Raps-Transformation wurde das Konstrukt aus dem Vektor PCR-Script als EcoRI-Fragment in den Vektor pZPNBN einkloniert. pZPNBN ist ein pPZP200 Derivat (Hajdukiewicz, P., et al., (1994) PMB 25(6): 989-94), dem zuvor eine Phosphinotricinresistenz unter der Kontrolle des  
35 NOS-Promotors vor dem NOS-Terminator eingefügt worden war. (Figur 7, Konstrukt XI)

#### Beispiel 7: Klonierung eines genomischen Fragments der Fumarylacetoacetat Isomerase aus *Arabidopsis thaliana*

40

- Mittels der Protein-Sequenz der FAAH aus *Emericella nidulans* wurde ein Blast-Search durchgeführt und aus *A. thaliana* eine Proteinsequenz identifiziert, die zu 59 % Homologie aufwies. FAAH aus *A. thaliana* hat die Accessionnummer AC002131. Mittels der ID-  
45 Nummer der Proteinsequenz konnte die DNA-Sequenz ermittelt werden und Oligonukleotide abgeleitet werden.

Dem 5'-Oligonukleotid wurde eine SalI und dem 3'-Oligonukleotid eine Asp718 Restriktionsschnittstelle angefügt. Das Oligonukleotid am 5'-Ende von FAAH umfasst die Sequenz

5 5'-atgtcgacGGAACTCTGAACCATAT-3' (SEQ ID NO: 31)

beginnend mit Base 40258 des BAC F12F1, das Oligonukleotid am 3'Ende umfasst die Sequenz:

10 5'-atggtaccGAATGTGATGCCTAAGT-3' (SEQ ID NO: 32)

beginnend mit dem Basenpaar 39653 des BACs. Die PCR-Reaktion wurde mit der Taq Polymerase (Hersteller: TaKaRa Shuzo Co., Ltd.) durchgeführt. Der Ansatz hatte folgende Zusammensetzung: 10 µl

15 Puffer (20 mM Tris-HCl pH 8,0, 100 mM KCl, 0,1 mM EDTA, 1 mM DTT, 0,5 % Tween20, 0,5 % Nonidet P-40, 50 % Glycerol), jeweils 100pmol der beiden Oligonukleotide, jeweils 20 nM an dATP, dCTP, dGTP, dTTP, 2,5 Einheiten Taq-Polymerase, 1 µg genomische DNA, destilliertes Wasser add 100 µl. Das PCR-Programm lautete:

20

- 5 Zyklen mit: 94°C (4 sec), 52°C (30 sec), 72°C (1 min)
- 5 Zyklen mit: 94°C (4 sec), 50°C (30 sec), 72°C (1 min)
- 25 Zyklen mit: 94°C (4 sec), 48°C (30 sec), 72°C (1 min)

25 Das Fragment ((SEQ ID NO: 42) wurde mittels Nucleo-Spin Extract (Machery-Nagel) gereinigt und nach Herstellerangaben in den Vektor pGEMTeasy von Promega kloniert (Figur 8, Konstrukt XII).

Die Richtigkeit des Fragments wurde durch Sequenzierung überprüft. Mittels der durch die Primer an die Sequenz angefügten  
30 Restriktionsschnittstellen wurde das Gen als SalI/Asp718-Fragment in den entsprechend geschnittenen Vektor BinAR (Höfgen, R. und Willmitzer, L., Plant Sci. 66: 221-230, 1990) kloniert. Dieser enthält den 35S-Promotor des Blumenkohlmosaikvirus und die OCS-  
35 Terminationssequenz (Fig. 9, Konstrukt XIII).

Zur Erstellung eines binären Vektors zur Arabidopsis- und Raps-Transformation wurde das Konstrukt aus dem Vektor pBinAR als EcoRI/HindIII-Fragment in den Vektor pZPNBN einkloniert. pZPNBN  
40 ist ein pZP200 Derivat (Hajdukewicz, P. et al., Plant Molecular Biology, 25: 989-994, 1994), dem zuvor eine Phosphinotricin-resistenz unter der Kontrolle des NOS-Promotors vor dem NOS-Terminator eingefügt worden war (Figur 9, Konstrukt XIV).



## Beispiel 8:

Herstellung transgener *Arabidopsis thaliana* Pflanzen

Wildtyp *Arabidopsis thaliana* Pflanzen (cv. Columbia) wurden mit  
5 dem *Agrobacterium tumefaciens* Stamm (EHA105) auf Grundlage einer  
modifizierten Methode der Vacuum Infiltrationsmethode nach Clough  
und Bent (Clough, S. and Bent A., Plant J. 16(6):735-43, 1998) und  
nach Bechtold, et al. (Bechtold, N., et al., CR Acad Sci Paris.  
1144(2):204-212, 1993) transformiert. Die verwendeten *Agrobacte-*  
10 *rium tumefaciens* Zellen waren zuvor mit den Plasmiden pZPNBN-  
MAA1anti bzw pZPNBN-FAA1anti transformiert worden.

Samen der Primärtransformanten wurden auf Grundlage der Phosphi-  
notricinresistenz geselektiert, indem Saatgut ausgelegt wurde und  
15 die Keimlinge mit dem Herbizid Basta (Phosphinotricin) besprüht  
wurden. Basta resistente Keimlinge wurden vereinzelt und als  
vollentwickelte Pflanzen zur biochemischen Analyse verwendet.

Beispiel 9: Herstellung transgener Raps (*Brassica napus*) Pflanzen

20

Die Herstellung transgener Raps Pflanzen orientierte sich an  
einem Protokoll von Bade, J.B. und Damm, B. (Bade, J.B. und Damm,  
B. (1995) in: Gene Transfer to Plants, Potrykus, I. und Spangen-  
berg, G., eds, Springer Lab Manual, Springer Verlag, 1995,  
25 30-38), in welchem auch die Zusammensetzung der verwendeten  
Medien und Puffer angegeben ist.

Die Transformation erfolgte mit dem *Agrobacterium tumefaciens*  
Stamm EHA105. Zur Transformation wurde entweder das Plasmid  
30 pPTVHPPDopHGDanti (Figur 4, Konstrukt VI) oder nach Anzucht ge-  
mischte Kulturen von *Agrobakterien* mit den Plasmiden pPTVHGDanti  
(Figur 4, Konstrukt V) und pPZP200HPPDop (Figur 5, Konstrukt VII)  
verwendet. Samen von *Brassica napus* var. Westar wurden mit 70 %  
Ethanol (v/v) oberflächensteril gemacht, 10 Minuten bei 55°C in  
35 Wasser gewaschen, in 1 %iger Hypochlorit-Lösung (25 % v/v Tee-  
pol, 0,1 % v/v Tween 20) für 20 Minuten inkubiert und sechsmal mit  
sterilem Wasser für jeweils 20 Minuten gewaschen. Die Samen wur-  
den drei Tage auf Filterpapier getrocknet und 10-15 Samen in  
einem Glaskolben mit 15 ml Keimungsmedium zur Keimung gebracht.  
40 Von mehreren Keimlingen (ca. 10 cm gross) wurden die Wurzeln und  
Apices entfernt und die verbleibenden Hypokotyle in ca. 6 mm  
lange Stücke geschnitten. Die so gewonnenen ca. 600 Explantate  
wurden 30 Minuten mit 50 ml Basalmedium gewaschen und in einem  
300 ml Kolben überführt. Nach Zugabe von 100 ml Kallusinduktions-  
45 medium wurden die Kulturen für 24 Stunden bei 100 U/min  
inkubiert.

- Von den *Agrobacterium* Stämmen wurden Übernachtskulturen bei 29°C in Luria Broth-Medium mit Kanamycin (20 mg/l) angesetzt, davon 2 ml in 50 ml Luria Broth-Medium ohne Kanamycin für 4 Stunden bei 29°C bis zu einer OD<sub>600</sub> von 0,4-0,5 inkubiert. Nach der Pelletierung
- 5 der Kultur bei 2000 U/min für 25 min wurde das Zellpellet in 25 ml Basalmedium resuspendiert. Die Konzentration der Bakterien in der Lösung wurde durch Zugabe von weiterem Basalmedium auf eine OD<sub>600</sub> von 0,3 eingestellt. Zur Kotransformation wurde die Lösung der beiden Stämme zu gleichen Teilen vermischt.
- 10 Aus den Raps-Explanten wurde das Kallus-Induktionsmedium mit sterilen Pipetten entfernt, 50 ml *Agrobacterium*-Lösung hinzugefügt, vorsichtig gemischt und für 20 min inkubiert. Die *Agrobacterien*-Suspension wurde entfernt, die Raps-Explantate für 1 min mit
- 15 50 ml Kallus-Induktionsmedium gewaschen und anschliessend 100 ml Kallus-Induktionsmedium hinzugefügt. Die Co-Kultivierung wurde für 24 h auf einem Rotationsschüttler bei 100 U/min durchgeführt. Die Co-Kultivierung wurde durch Wegnahme des Kallus-Induktionsmediums gestoppt und die Explantate zweimal für jeweils 1 min mit
- 20 25 ml und zweimal für 60 min mit jeweils 100 ml Waschmedium bei 100 U/min gewaschen. Das Waschmedium mit den Explantaten wurde in 15 cm Petrischalen überführt und das Medium mit sterilen Pipetten entfernt.
- 25 Zur Regeneration wurden jeweils 20-30 Explantate in 90 mm Petrischalen überführt, welche 25 ml Spross-Induktionsmedium mit Phosphinotricin enthielten. Die Petrischalen wurden mit 2 Lagen Leukopor verschlossen und bei 25°C und 2000 lux bei Photoperioden von 16 Stunden Licht / 8 Stunden Dunkelheit inkubiert. Alle 12
- 30 Tage wurden die sich entwickelnden Kalli auf frische Petrischalen mit Spross-Induktionsmedium umgesetzt. Alle weiteren Schritte zur Regeneration ganzer Pflanzen wurde wie von Bade, J.B und Damm, B. (in: Gene Transfer to Plants, Potrykus, I. und Spangenberg, G., eds, Springer Lab Manual, Springer Verlag, 1995, 30-38) beschrieben durchgeführt.
- 35

#### Beispiel 10:

##### Untersuchung der transgenen Pflanzen

- 40 Um zu bestätigen, dass durch die Inhibition der HGD-, MAAI,- und/oder FAAH die Vitamin E Biosynthese in den transgenen Pflanzen beeinflusst wird, werden die Tocopherol- und Tocotrienol-Gehalte in Blätter und Samen der mit den beschriebenen Konstrukten transformierten Pflanzen (*Arabidopsis thaliana*, *Brassica napus*) analysiert. Dazu werden die transgenen Pflanzen im Gewächshaus kultiviert und Pflanzen, die die antisense RNA von der HGD-, MAAI,- und/oder FAAH exprimieren, mittels einer Northern-Blot Analyse
- 45

untersucht. In Blättern und Samen dieser Pflanzen wird der Tocopherolgehalt und der Tocotrienolgehalt ermittelt. Der Aufschluß des Pflanzenmaterials erfolgte durch dreimalige Inkubation im Ependorfschüttler bei 30°C, 1000rpm in 100 % Methanol für 15 Minuten, wobei die jeweils erhaltenen Überstände vereinigt wurden. Weitere Inkubationsschritte ergaben keine weitere Freisetzung von Tocopherolen oder Tocotrienolen. Um Oxidation zu vermeiden, wurden die erhaltenen Extrakte direkt nach der Extraktion mit Hilfe einer Waters Alliance 2690 HPLC Anlage analysiert. Tocopherole und Tocotrienole wurden über eine reverse Phase Säule( ProntoSil 200-3-C30, Bischoff) mit einer mobilen Phase von 100 % Methanol getrennt und anhand von Standards (Merck) identifiziert. Als Detektionssystem diente die Fluoreszenz der Substanzen (Anregung 295nm, Emission 320 nm) die mit Hilfe eines Jasco Fluoreszenzdetektors FP 920 nachgewiesen wurde.

In allen Fällen ist die Tocopherol- bzw. Tocotrienol-Konzentration in transgenen Pflanzen, die zusätzlich eine erfindungsgemäße Nukleinsäure exprimieren, im Vergleich zu nicht transformierten Pflanzen erhöht.

25

30

35

40

45

## Patentansprüche

1. Verfahren zur Bildung von Vitamin E durch Beeinflussung der  
5 Vitamin E-Biosynthese, dadurch gekennzeichnet, dass man den  
Homogentisatabbau durch Verminderung der Homogentisat-1,2-  
dioxygenase (HGD) - Aktivität, Maleylacetocetatisomerase  
(MAAI) - Aktivität und/oder Fumarylacetoacetathydrolase  
(FAAH) - Aktivität reduziert.
- 10 2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass man  
die MAAI-Aktivität und/oder die FAAH-Aktivität reduziert und  
gleichzeitig
- 15 a) die Umsetzung von Homogentisat zu Vitamin E verbessert,  
oder
- b) die Biosynthese von Homogentisat verbessert.
- 20 3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass man  
die HGD-Aktivität reduziert und gleichzeitig
- a) die Umsetzung von Homogentisat zu Vitamin E verbessert,  
oder
- 25 b) das TyrA-Gen überexprimiert.
4. Verfahren zur vermehrten Bildung von Vitamin E durch Beein-  
flussung der Vitamin E-Biosynthese, dadurch gekennzeichnet,  
30 dass man
- a) die Umsetzung von Homogentisat zu Vitamin E, und gleich-  
zeitig
- 35 b) die Biosynthese von Homogentisat  
verbessert.
5. Verfahren nach Anspruch 1 bis 3, dadurch gekennzeichnet, dass  
40 man die Kultur eines pflanzlichen Organismus mit Inhibitoren  
der MAAI, HGD oder FAAH behandelt.

6. Nukleinsäurekonstrukt, enthaltend eine Nukleinsäuresequenz (anti-MAAI/FAAH), welche zu einer Reduktion der MAAI-Aktivität oder FAAH-Aktivität befähigt ist, oder ein funktionales Äquivalent davon.
- 5
7. Nukleinsäurekonstrukt nach Anspruch 6, enthaltend zusätzlich
- 10 a) eine Nukleinsäuresequenz (pro-HG), welche zu einer Steigerung der Homogentisat (HG) Biosynthese befähigt ist, oder ein funktionales Äquivalent davon; oder
- 15 b) eine Nukleinsäuresequenz (pro-VitaminE), welche zu einer Steigerung der Vitamin E Biosynthese ausgehend vom Homogentisat befähigt ist, oder ein funktionales Äquivalent davon; oder
- c) eine Kombination von a) und b).
8. Nukleinsäurekonstrukt, enthaltend eine Nukleinsäuresequenz (anti-HGD), welche zu einer Inhibition der HGD befähigt ist, oder für ein funktionales Äquivalent davon.
- 20
9. Nukleinsäurekonstrukt nach Anspruch 8, enthaltend zusätzlich
- 25 a) eine Nukleinsäuresequenz kodierend für bifunktionale Chorismatmutase-Prephenatdehydrogenase Enzyme (TyrA) oder ein funktionales Äquivalent davon; oder
- 30 b) eine Nukleinsäuresequenz (pro-VitaminE), welche zu einer Steigerung der Vitamin E Biosynthese ausgehend vom Homogentisat befähigt ist, oder ein funktionales Äquivalent davon; oder
- 35 c) eine Kombination von a) und b).
10. Nukleinsäurekonstrukt, enthaltend eine Nukleinsäuresequenz (pro-HG), welche zu einer Steigerung der Homogentisat (HG) Biosynthese befähigt ist, oder ein funktionales Äquivalent davon, und gleichzeitig eine Nukleinsäuresequenz (pro-VitaminE), welche zu einer Steigerung der Vitamin E Biosynthese ausgehend vom Homogentisat befähigt ist, oder ein funktionales Äquivalent davon.
- 40
11. Nukleinsäurekonstrukt nach Anspruch 6 bis 10 enthaltend eine
- 45 anti-MAAI/FAAH-Sequenz bzw. anti-HGD-Sequenz, die

- a) zu einer antisense-Nukleinsäuresequenz transkribierbar ist, welche zur Inhibition der MAAI/ FAAH-Aktivität bzw. der HGD-Aktivität befähigt ist, oder
- 5 b) eine Inaktivierung der MAAI/FAAH bzw. HGD durch eine homologe Rekombination bewirkt, oder
- c) für einen Bindungsfaktor kodiert, der an die Gene der MAAI/FAAH bzw. HGD bindet und so die Transkription dieser
- 10 Gene vermindert.
12. Nukleinsäurekonstrukt nach Anspruch 7 und 10 enthaltend eine proHG-Sequenz ausgewählt aus den Genen kodierend für eine HPPD, TyrA.
- 15 13. Nukleinsäurekonstrukt nach Anspruch 7, 9 und 10 enthaltend eine proVitaminE-Sequenz ausgewählt aus den Genen kodierend für eine HPGT, Geranylgeranyloxidoreduktase, 2-Methyl-6-phytylplastoquinol-methyltransferase,  $\gamma$ -Tocopherol-methyltransferase.
- 20 14. Rekombinanter Vektor, enthaltend
- a) ein Nukleinsäurekonstrukt nach einem der Ansprüche 6 bis
- 25 13; oder
- b) eine Nukleinsäure, die für eine HGD, MAAH oder FAAH kodiert, sowie funktionelle Äquivalente davon, oder
- 30 c) eine Kombination der Möglichkeiten a) und b).
15. Rekombinanter Vektor nach den Anspruch 14, wobei die Nukleinsäuren oder Nukleinsäurekonstrukte funktionell mit einer genetischen Kontrollsequenz verbunden sind und der die Fähigkeit zur Transkription von sense oder antisense-RNA hat.
- 35 16. Transgener Organismus, transformiert mit einem Nukleinsäurekonstrukt gemäss den Ansprüchen 6 bis 13 oder einem rekombinanten Vektor gemäss den Ansprüchen 14 oder 15.
- 40 17. Transgener Organismus nach Anspruch 16 ausgewählt aus Bakterien, Hefen, Pilzen, Moosen, tierischen und pflanzlichen Organismen.

18. Zellkulturen, Teile, transgenes Vermehrungsgut, oder Früchte abgeleitet von einem transgenen Organismus nach den Ansprüchen 16 oder 17.
- 5 19. Verwendung eines transgenen Organismus nach einem der Ansprüche 16 oder 17 oder von diesem abgeleitete Zellkulturen, Teile, transgenes Vermehrungsgut oder Früchte nach Anspruch 18 als Nahrungs- oder Futtermittel oder zur Isolation von Vitamin E.
- 10 20. Antikörper, proteinbindende oder DNA-bindende Faktoren gegen Polypeptide mit HGD-, MAAI- oder FAAH-Aktivität, deren Gene oder cDNAs.
- 15 21. Verwendung von Polypeptiden mit HGD-, MAAI- oder FAAH-Aktivität, deren Gene oder cDNAs zum Auffinden von Inhibitoren der HGD, MAAI oder FAAH.
- 20 22. Verfahren zum Auffinden von Inhibitoren der MAAI, HGD oder FAAH, dadurch gekennzeichnet, daß man die enzymatische Aktivität der MAAI, HGD oder FAAH in Gegenwart einer chemischen Verbindung misst und bei Erniedrigung der enzymatischen Aktivität im Vergleich zur nicht gehemmten Aktivität die chemische Verbindung einen Inhibitor darstellt.
- 25 23. Verwendung von Inhibitoren der HGD, MAAI oder FAAH, erhältlich gemäß einem Verfahren nach Anspruch 22, als Wachstumsregulatoren.

30

35

40

45

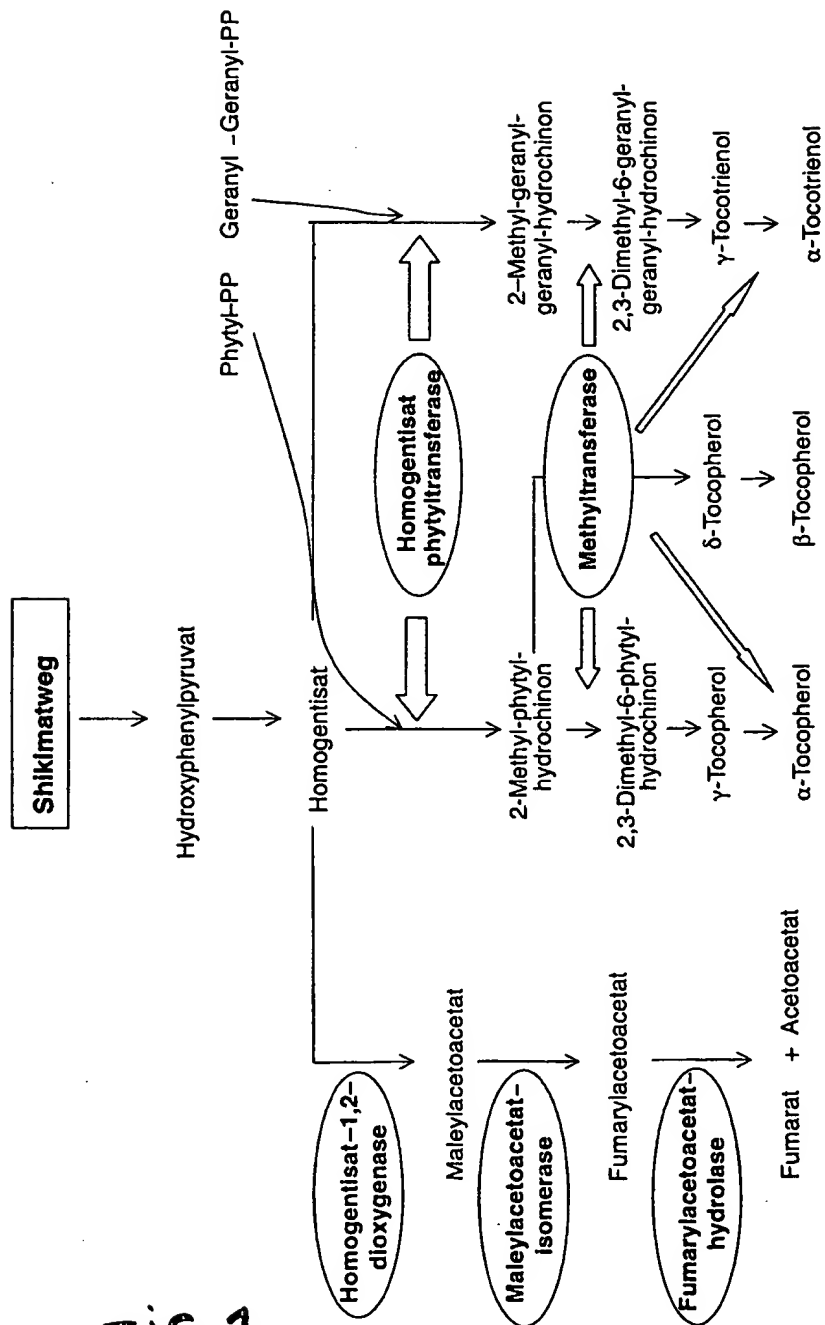


FIG. 1



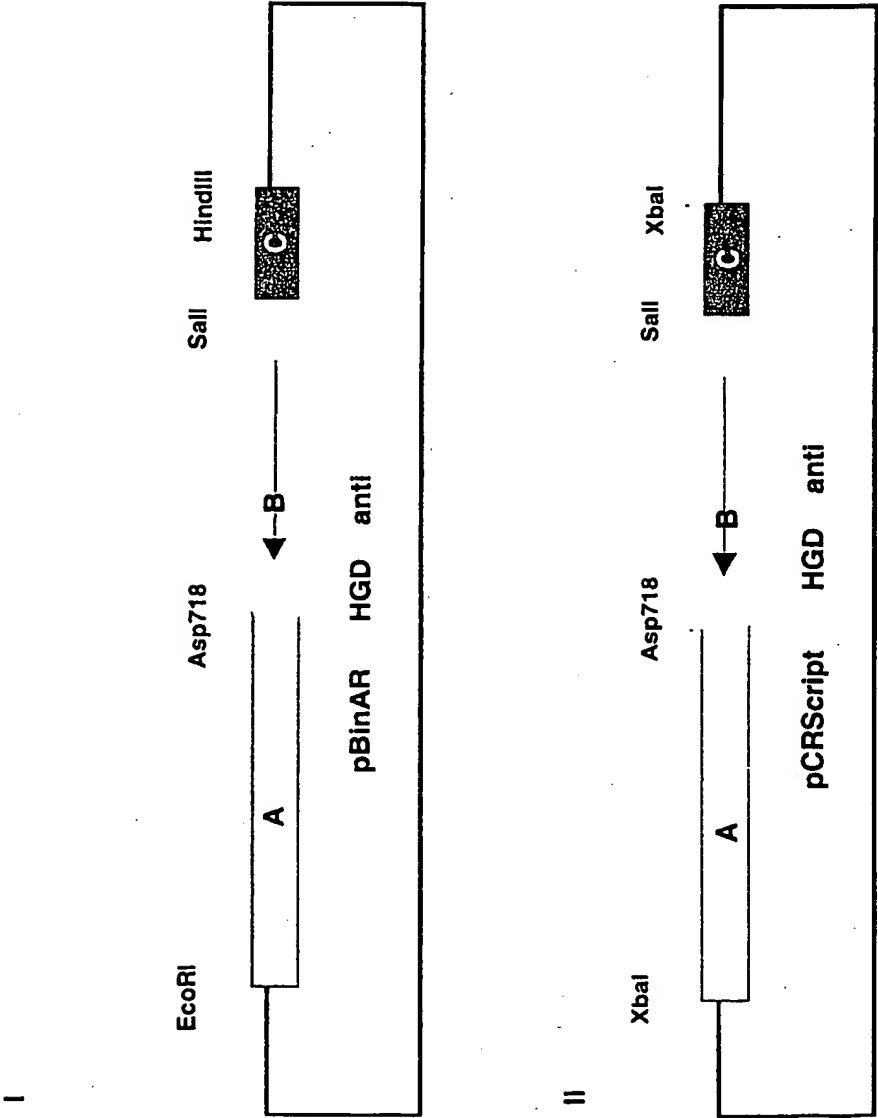


FIG. 2

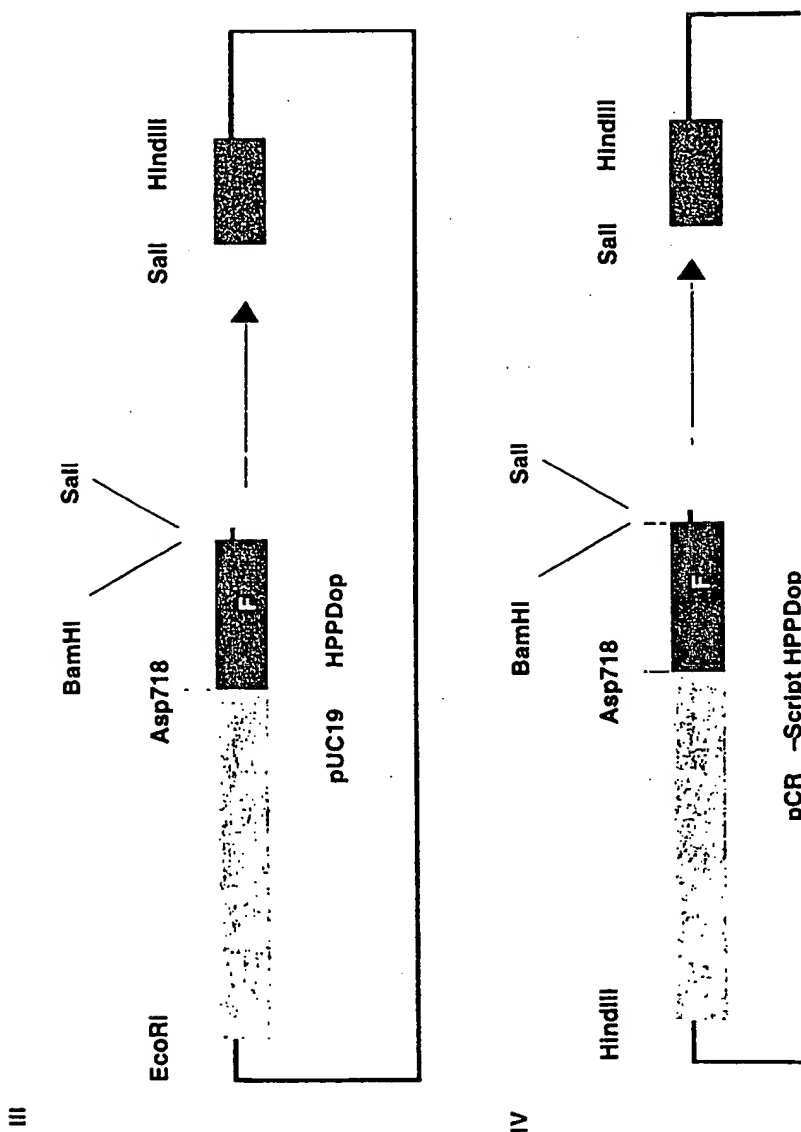


FIG. 3

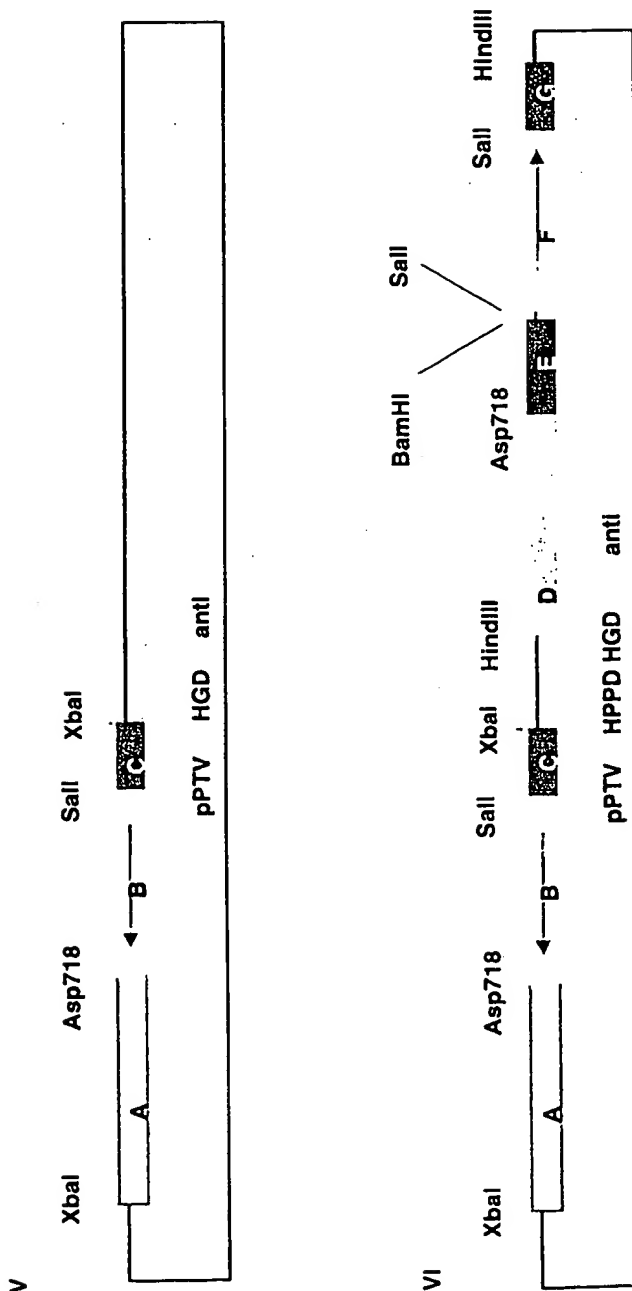


Fig. 4

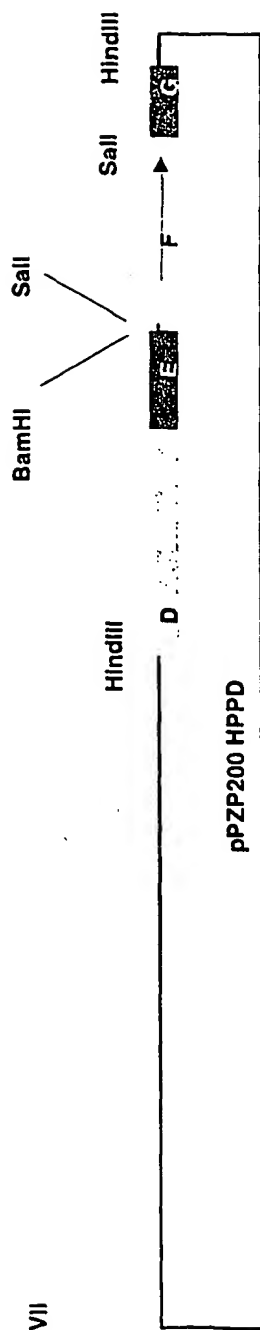
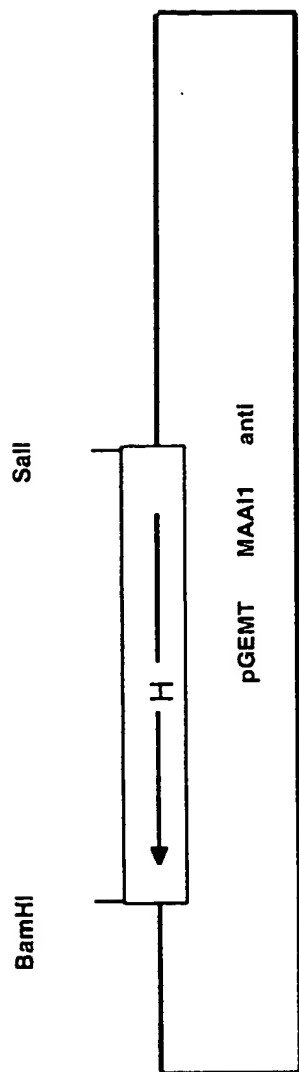
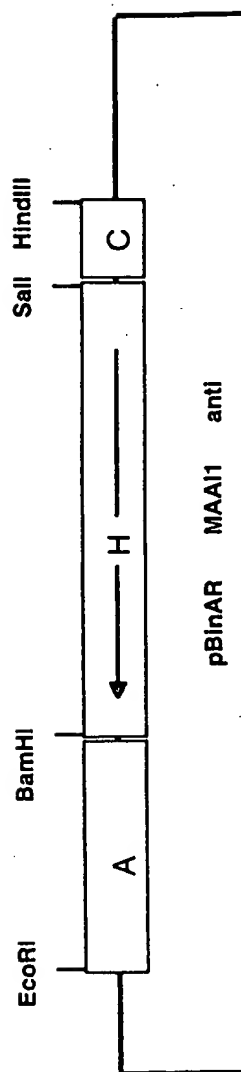


FIG. 5

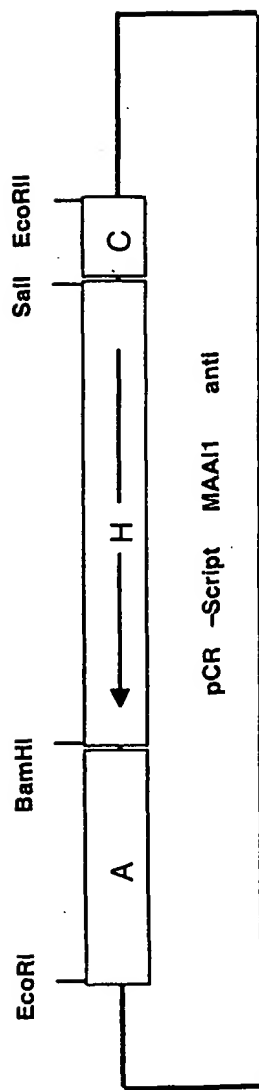


VIII

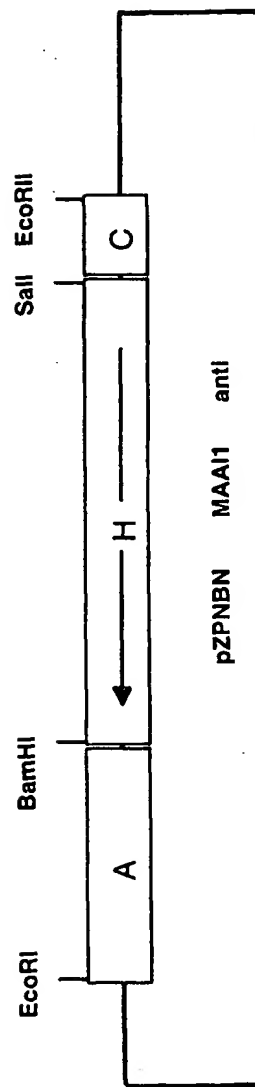


IX

FIG. 6



X



XI

FIG. 7

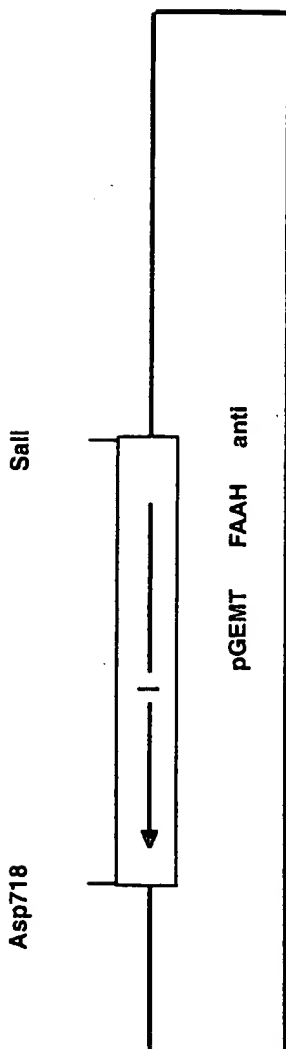
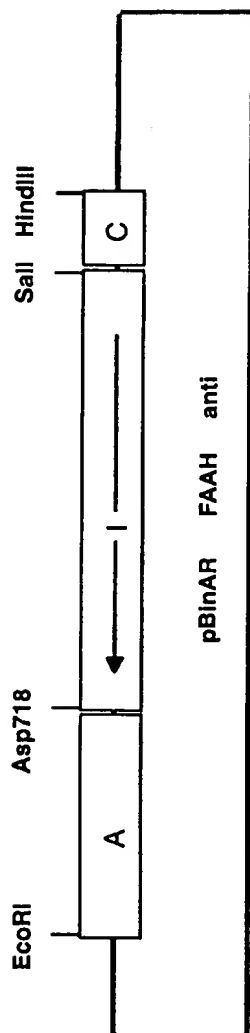
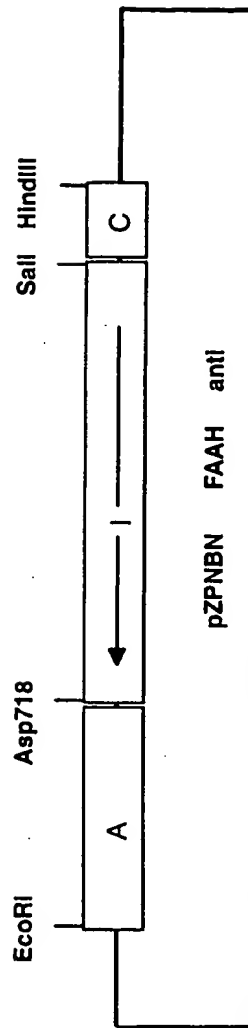


FIG. 8



XIII



XIV

FIG. 9



## SEQUENZPROTOKOLL

&lt;110&gt; SunGene GmbH &amp; Co. KGaA

&lt;120&gt; Verbesserte Verfahren zur Vitamin E Biosynthese

&lt;130&gt; NAE445/2000

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 44

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 2151

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; gene

&lt;222&gt; (1)..(2151)

&lt;223&gt; gene for homogentisate-1,2-dioxygenase (HGD)

&lt;400&gt; 1

```
atggaagaga agaagaagga gcttgaagag ttgaagtatc aatcagggtt tggttaaccac 60
ttctcatcgg aagcaatcgc cggagcttta ccggttagatc agaacagtc tcttctttgt 120
ccttacgggc tttacgccga acagatctcc ggtacttctt tcacttctcc tcgcaagctc 180
aatcaaagaa ggtacatcat catttcaatt gtaagttttg gataatttcg ttgaattgat 240
tgatcttcat cttgtttttt ttttcagttg gttgtaccgg gttaaaccat cggttacaca 300
tgaaccgttc aagcctcgtg taccagctca taagaagctt gtgagtgagt ttgatgcac 360
aaatagtcgt acgaatccga ctcagcttcg gtggagacct gaggatatc ctgattcggg 420
gattgatttc gttgatgggt tatttaccat ttgtggagct ggaagctcgt tcttctcgcc 480
tggcttcgct attcacatgt aaaaaactct tctttttatt ttggtatctt tgggtgtagat 540
cagtataca taaagtaatg atcttttgta ttcattttgt tttgaaggta tgtggctaac 600
acaggaatga aagactccgc attttgcaac gctgatgggt acttcttggt agttcctcaa 660
acaggaagta agttagtagt cccaatgcct taccctacca catctttggg aaataaagtc 720
agtcattgat tgagaatgga ttcaagatag tcttgatca gttctgatag tttgagtggg 780
tgttttaggg ctatggattg aaactgagtg tgggaaggctt ttggttaactc ctggtgagat 840
tgctgttata ccacaagggt tccgtttctc catagattta ccggatggga agtctcgtgg 900
ttatgttgct gaaatctatg gggctcattt tcagcttctt gatcttgga caataggtag 960
tcttgagttc ttttagattc agccggaata acatggattc tccgcaagaa tcttattggg 1020
ggatgtggac aggtgctaag ggtcttgctg catcaagaga ttttcttgca ccaacagcat 1080
ggtttgagga tggattgcgg cctgaatata caattgttca gaagtttggc ggtgaactct 1140
ttactgctaa acaagatttc tctccattca atgtggttgc ctggcatggc aattacgtgc 1200
cttataagggt gagtacattg tttattgagc ctaatcttgt aaaacgttaa tgcattgttt 1260
ttctgagaat ttcaatttct gtctgcagta tgacctgaag aagttctgtc catacaaac 1320
tgtgtcttta gatcatggag atccatctat aaatacaggt tgggtggtcat ctgcgctaaa 1380
tcgattcttc tttttgtttt gttatgggtt ggttacttgt tctttattgt aatcacactc 1440
tttgggtgaa ttattgtact ctcagtcctt acagcaccaa ctgataaacc tgggtgtggc 1500
ttgcttgatt ttgtcatatt tctctctcga tgggtgggtg ctgagcatac tttctgacct 1560
ccttactatc atcgtaactg catgagtga tttatgggct taatctacgg tgcatacgag 1620
```

## 2

gtaagctgct tgaagttcct gcttctgcaa atcattagct ggcttggtgt atcctcctac 1680  
 tgaaatctgt aaactgactc caccattcac aggcgaaagc tgatggattt ctccctggcg 1740  
 gtgcaagtct tcatagctgt atgacacctc atgggtccaga tactaccacg tacgaggtat 1800  
 caatccatct tatgcacagc agcaactaca cgtttgattt cattttcctc cgagatcatg 1860  
 tctaaatcta acccctgaat gtaaaattaa gtctgaagca tttttataat tgtttttag 1920  
 gcgacaattg ctcgagtaaa tgcaatggct ccttctaaac tcacaggtac gatggccttc 1980  
 atgttcgaat cagcattgat ccctagagtc tgtcattggg ctctggagtc tcctttcctg 2040  
 gatcacgact actaccagtg ttggattggc ctcaagtctc atttctcgcg cataagcttg 2100  
 gacaagacaa atgttgaatc aacagagaaa gaaccaggag cttcggagta a 2151

<210> 2

<211> 1386

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1383)

<223> cDNA coding for homogentisate-1,2-dioxygenase  
 (HGD)

<400> 2

atg gaa gag aag aag aag gag ctt gaa gag ttg aag tat caa tca ggt	48
Met Glu Glu Lys Lys Lys Glu Leu Glu Glu Leu Lys Tyr Gln Ser Gly	
1 5 10 15	
ttt ggt aac cac ttc tca tcg gaa gca atc gcc gga gct tta ccg tta	96
Phe Gly Asn His Phe Ser Ser Glu Ala Ile Ala Gly Ala Leu Pro Leu	
20 25 30	
gat cag aac agt cct ctt ctt tgt cct tac ggt ctt tac gcc gaa cag	144
Asp Gln Asn Ser Pro Leu Leu Cys Pro Tyr Gly Leu Tyr Ala Glu Gln	
35 40 45	
atc tcc ggt act tct ttc act tct cct cgc aag ctc aat caa aga agt	192
Ile Ser Gly Thr Ser Phe Thr Ser Pro Arg Lys Leu Asn Gln Arg Ser	
50 55 60	
tgg ttg tac cgg gtt aaa cca tcg gtt aca cat gaa ccg ttc aag cct	240
Trp Leu Tyr Arg Val Lys Pro Ser Val Thr His Glu Pro Phe Lys Pro	
65 70 75 80	
cgt gta cca gct cat aag aag ctt gtg agt gag ttt gat gca tca aat	288
Arg Val Pro Ala His Lys Lys Leu Val Ser Glu Phe Asp Ala Ser Asn	
85 90 95	
agt cgt acg aat ccg act cag ctt cgg tgg aga cct gag gat att cct	336
Ser Arg Thr Asn Pro Thr Gln Leu Arg Trp Arg Pro Glu Asp Ile Pro	
100 105 110	
gat tcg gag att gat ttc gtt gat ggg tta ttt acc att tgt gga gct	384
Asp Ser Glu Ile Asp Phe Val Asp Gly Leu Phe Thr Ile Cys Gly Ala	
115 120 125	

## 3

gga agc tcg ttt ctt cgc cat ggc ttc gct att cac atg tat gtg gct	432
Gly Ser Ser Phe Leu Arg His Gly Phe Ala Ile His Met Tyr Val Ala	
130 135 140	
aac aca gga atg aaa gac tcc gca ttt tgc aac gct gat ggt gac ttc	480
Asn Thr Gly Met Lys Asp Ser Ala Phe Cys Asn Ala Asp Gly Asp Phe	
145 150 155 160	
ttg tta gtt cct caa aca gga agg cta tgg att gaa act gag tgt gga	528
Leu Leu Val Pro Gln Thr Gly Arg Leu Trp Ile Glu Thr Glu Cys Gly	
165 170 175	
agg ctt ttg gta act cct ggt gag att gct gtt ata cca caa ggt ttc	576
Arg Leu Leu Val Thr Pro Gly Glu Ile Ala Val Ile Pro Gln Gly Phe	
180 185 190	
cgt ttc tcc ata gat tta ccg gat ggg aag tct cgt ggt tat gtt gct	624
Arg Phe Ser Ile Asp Leu Pro Asp Gly Lys Ser Arg Gly Tyr Val Ala	
195 200 205	
gaa atc tat ggg gct cat ttt cag ctt cct gat ctt gga cca ata ggt	672
Glu Ile Tyr Gly Ala His Phe Gln Leu Pro Asp Leu Gly Pro Ile Gly	
210 215 220	
gct aat ggt ctt gct gca tca aga gat ttt ctt gca cca aca gca tgg	720
Ala Asn Gly Leu Ala Ala Ser Arg Asp Phe Leu Ala Pro Thr Ala Trp	
225 230 235 240	
ttt gag gat gga ttg cgg cct gaa tac aca att gtt cag aag ttt ggc	768
Phe Glu Asp Gly Leu Arg Pro Glu Tyr Thr Ile Val Gln Lys Phe Gly	
245 250 255	
ggt gaa ctc ttt act gct aaa caa gat ttc tct cca ttc aat gtg gtt	816
Gly Glu Leu Phe Thr Ala Lys Gln Asp Phe Ser Pro Phe Asn Val Val	
260 265 270	
gcc tgg cat ggc aat tac gtg cct tat aag tat gac ctg aag aag ttc	864
Ala Trp His Gly Asn Tyr Val Pro Tyr Lys Tyr Asp Leu Lys Lys Phe	
275 280 285	
tgt cca tac aac act gtg ctt tta gat cat gga gat cca tct ata aat	912
Cys Pro Tyr Asn Thr Val Leu Leu Asp His Gly Asp Pro Ser Ile Asn	
290 295 300	
aca gtc ctt aca gca cca act gat aaa cct ggt gtg gcc ttg ctt gat	960
Thr Val Leu Thr Ala Pro Thr Asp Lys Pro Gly Val Ala Leu Leu Asp	
305 310 315 320	
ttt gtc ata ttt cct cct cga tgg ttg gtt gct gag cat act ttt cga	1008
Phe Val Ile Phe Pro Pro Arg Trp Leu Val Ala Glu His Thr Phe Arg	
325 330 335	

4

cct cct tac tat cat cgt aac tgc atg agt gaa ttt atg ggc tta atc 1056  
 Pro Pro Tyr Tyr His Arg Asn Cys Met Ser Glu Phe Met Gly Leu Ile  
 340 345 350

tac ggt gca tac gag gcg aaa gct gat gga ttt ctc cct ggc ggt gca 1104  
 Tyr Gly Ala Tyr Glu Ala Lys Ala Asp Gly Phe Leu Pro Gly Gly Ala  
 355 360 365

agt ctt cat agc tgt atg aca cct cat ggt cca gat act acc acg tac 1152  
 Ser Leu His Ser Cys Met Thr Pro His Gly Pro Asp Thr Thr Thr Tyr  
 370 375 380

gag gcg aca att gct cga gta aat gca atg gct cct tct aaa ctc aca 1200  
 Glu Ala Thr Ile Ala Arg Val Asn Ala Met Ala Pro Ser Lys Leu Thr  
 385 390 395 400

ggt acg atg gct ttc atg ttc gaa tca gca ttg atc cct aga gtc tgt 1248  
 Gly Thr Met Ala Phe Met Phe Glu Ser Ala Leu Ile Pro Arg Val Cys  
 405 410 415

cat tgg gct ctg gag tct cct ttc ctg gat cac gac tac tac cag tgt 1296  
 His Trp Ala Leu Glu Ser Pro Phe Leu Asp His Asp Tyr Tyr Gln Cys  
 420 425 430

tgg att ggc ctc aag tct cat ttc tcg cgc ata agc ttg gac aag aca 1344  
 Trp Ile Gly Leu Lys Ser His Phe Ser Arg Ile Ser Leu Asp Lys Thr  
 435 440 445

aat gtt gaa tca aca gag aaa gaa cca gga gct tcg gag taa 1386  
 Asn Val Glu Ser Thr Glu Lys Glu Pro Gly Ala Ser Glu  
 450 455 460

&lt;210&gt; 3

&lt;211&gt; 461

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 3

Met Glu Glu Lys Lys Lys Glu Leu Glu Glu Leu Lys Tyr Gln Ser Gly  
 1 5 10 15

Phe Gly Asn His Phe Ser Ser Glu Ala Ile Ala Gly Ala Leu Pro Leu  
 20 25 30

Asp Gln Asn Ser Pro Leu Leu Cys Pro Tyr Gly Leu Tyr Ala Glu Gln  
 35 40 45

Ile Ser Gly Thr Ser Phe Thr Ser Pro Arg Lys Leu Asn Gln Arg Ser  
 50 55 60

Trp Leu Tyr Arg Val Lys Pro Ser Val Thr His Glu Pro Phe Lys Pro  
 65 70 75 80

Arg Val Pro Ala His Lys Lys Leu Val Ser Glu Phe Asp Ala Ser Asn  
 85 90 95

## 5

Ser Arg Thr Asn Pro Thr Gln Leu Arg Trp Arg Pro Glu Asp Ile Pro  
 100 105 110  
 Asp Ser Glu Ile Asp Phe Val Asp Gly Leu Phe Thr Ile Cys Gly Ala  
 115 120 125  
 Gly Ser Ser Phe Leu Arg His Gly Phe Ala Ile His Met Tyr Val Ala  
 130 135 140  
 Asn Thr Gly Met Lys Asp Ser Ala Phe Cys Asn Ala Asp Gly Asp Phe  
 145 150 155 160  
 Leu Leu Val Pro Gln Thr Gly Arg Leu Trp Ile Glu Thr Glu Cys Gly  
 165 170 175  
 Arg Leu Leu Val Thr Pro Gly Glu Ile Ala Val Ile Pro Gln Gly Phe  
 180 185 190  
 Arg Phe Ser Ile Asp Leu Pro Asp Gly Lys Ser Arg Gly Tyr Val Ala  
 195 200 205  
 Glu Ile Tyr Gly Ala His Phe Gln Leu Pro Asp Leu Gly Pro Ile Gly  
 210 215 220  
 Ala Asn Gly Leu Ala Ala Ser Arg Asp Phe Leu Ala Pro Thr Ala Trp  
 225 230 235 240  
 Phe Glu Asp Gly Leu Arg Pro Glu Tyr Thr Ile Val Gln Lys Phe Gly  
 245 250 255  
 Gly Glu Leu Phe Thr Ala Lys Gln Asp Phe Ser Pro Phe Asn Val Val  
 260 265 270  
 Ala Trp His Gly Asn Tyr Val Pro Tyr Lys Tyr Asp Leu Lys Lys Phe  
 275 280 285  
 Cys Pro Tyr Asn Thr Val Leu Leu Asp His Gly Asp Pro Ser Ile Asn  
 290 295 300  
 Thr Val Leu Thr Ala Pro Thr Asp Lys Pro Gly Val Ala Leu Leu Asp  
 305 310 315 320  
 Phe Val Ile Phe Pro Pro Arg Trp Leu Val Ala Glu His Thr Phe Arg  
 325 330 335  
 Pro Pro Tyr Tyr His Arg Asn Cys Met Ser Glu Phe Met Gly Leu Ile  
 340 345 350  
 Tyr Gly Ala Tyr Glu Ala Lys Ala Asp Gly Phe Leu Pro Gly Gly Ala  
 355 360 365  
 Ser Leu His Ser Cys Met Thr Pro His Gly Pro Asp Thr Thr Tyr  
 370 375 380  
 Glu Ala Thr Ile Ala Arg Val Asn Ala Met Ala Pro Ser Lys Leu Thr  
 385 390 395 400

6

Gly Thr Met Ala Phe Met Phe Glu Ser Ala Leu Ile Pro Arg Val Cys  
 405 410 415

His Trp Ala Leu Glu Ser Pro Phe Leu Asp His Asp Tyr Tyr Gln Cys  
 420 425 430

Trp Ile Gly Leu Lys Ser His Phe Ser Arg Ile Ser Leu Asp Lys Thr  
 435 440 445

Asn Val Glu Ser Thr Glu Lys Glu Pro Gly Ala Ser Glu  
 450 455 460

&lt;210&gt; 4

&lt;211&gt; 1227

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1224)

<223> cDNA coding for fumarylacetoacetate hydrolase  
 (FAAH)

&lt;400&gt; 4

atg gcg ttg ctg aag tct ttc atc gat gtt ggc tca gac tcg cac ttc 48  
 Met Ala Leu Leu Lys Ser Phe Ile Asp Val Gly Ser Asp Ser His Phe

1

5

10

15

cct atc cag aat ctc cct tat ggt gtc ttc aaa ccg gaa tcg aac tca 96  
 Pro Ile Gln Asn Leu Pro Tyr Gly Val Phe Lys Pro Glu Ser Asn Ser  
 20 25 30

act cct cgt cct gcc gtc gct atc ggc gat ttg gtt ctg gac ctc tcc 144  
 Thr Pro Arg Pro Ala Val Ala Ile Gly Asp Leu Val Leu Asp Leu Ser  
 35 40 45

gct atc tct gaa gct ggg ctt ttc gat ggt ctg atc ctt aag gac gca 192  
 Ala Ile Ser Glu Ala Gly Leu Phe Asp Gly Leu Ile Leu Lys Asp Ala  
 50 55 60

gat tgc ttt ctt cag cct aat ttg aat aag ttc ttg gcc atg gga cgg 240  
 Asp Cys Phe Leu Gln Pro Asn Leu Asn Lys Phe Leu Ala Met Gly Arg  
 65 70 75 80

cct gcg tgg aag gaa gcg cgt tct acg ctg caa aga atc ttg tca ttt 288  
 Pro Ala Trp Lys Glu Ala Arg Ser Thr Leu Gln Arg Ile Leu Ser Phe  
 85 90 95

ttg tta ttt ggc ttc aag gtt ttg gtt ttg gta tgt ttt cat gca gct 336  
 Leu Leu Phe Gly Phe Lys Val Leu Val Leu Val Cys Phe His Ala Ala  
 100 105 110

7

aat gaa cct atc ttg cga gac aat gat gtt ttg agg aga aaa tca ttc	384
Asn Glu Pro Ile Leu Arg Asp Asn Asp Val Leu Arg Arg Lys Ser Phe	
115 120 125	
cat cag atg agt aaa gtg gaa atg att gtt cct atg gtg att ggg gac	432
His Gln Met Ser Lys Val Glu Met Ile Val Pro Met Val Ile Gly Asp	
130 135 140	
tat aca gac ttc ttt gca tct atg cat cac gcg aag aac tgc gga ctt	480
Tyr Thr Asp Phe Phe Ala Ser Met His His Ala Lys Asn Cys Gly Leu	
145 150 155 160	
atg ttc cgt ggg cct gag aat gcg ata aac cca aat tgg ttt cgt ctt	528
Met Phe Arg Gly Pro Glu Asn Ala Ile Asn Pro Asn Trp Phe Arg Leu	
165 170 175	
ccc att gca tat cat gga cgg gca tca tct att gtc atc tct ggg act	576
Pro Ile Ala Tyr His Gly Arg Ala Ser Ser Ile Val Ile Ser Gly Thr	
180 185 190	
gac att att cga cca aga ggt cag ggc cat cca caa gga aac tct gaa	624
Asp Ile Ile Arg Pro Arg Gly Gln Gly His Pro Gln Gly Asn Ser Glu	
195 200 205	
cca tat ttt gga cct tcg aag aaa ctt gat ttt gag ctt gag atg gct	672
Pro Tyr Phe Gly Pro Ser Lys Lys Leu Asp Phe Glu Leu Glu Met Ala	
210 215 220	
gct gtg gtt ggt cca gga aat gaa ttg gga aag cct att gac gtg aat	720
Ala Val Val Gly Pro Gly Asn Glu Leu Gly Lys Pro Ile Asp Val Asn	
225 230 235 240	
aat gca gcc gat cat ata ttt ggt cta tta ctg atg aat gac tgg agt	768
Asn Ala Ala Asp His Ile Phe Gly Leu Leu Leu Met Asn Asp Trp Ser	
245 250 255	
gct agg gat att cag gcg tgg gag tat gta cct ctt ggt cct ttc ctg	816
Ala Arg Asp Ile Gln Ala Trp Glu Tyr Val Pro Leu Gly Pro Phe Leu	
260 265 270	
ggg aag agt ttt ggg act act ata tcc cct tgg att gtt acc ttg gat	864
Gly Lys Ser Phe Gly Thr Thr Ile Ser Pro Trp Ile Val Thr Leu Asp	
275 280 285	
gcg ctt gag cct ttt ggt tgt caa gct ccc aag cag gat cca cct cca	912
Ala Leu Glu Pro Phe Gly Cys Gln Ala Pro Lys Gln Asp Pro Pro Pro	
290 295 300	
ttg cca tat ttg gct gag aaa gag tct gta aat tac gat atc tcc ttg	960
Leu Pro Tyr Leu Ala Glu Lys Glu Ser Val Asn Tyr Asp Ile Ser Leu	
305 310 315 320	

8

gag cta gca cac cat acc gtt aac ggt tgc aat ttg agg cct ggt gat 1008  
 Glu Leu Ala His His Thr Val Asn Gly Cys Asn Leu Arg Pro Gly Asp  
                   325                                  330                                  335

ctc ctt ggc aca gga acc ata agc gga ccg gag cca gat tca tat ggg 1056  
 Leu Leu Gly Thr Gly Thr Ile Ser Gly Pro Glu Pro Asp Ser Tyr Gly  
                   340                                  345                                  350

tgc cta ctt gag ttg aca tgg aat gga cag aaa cct cta tca ctc aat 1104  
 Cys Leu Leu Glu Leu Thr Trp Asn Gly Gln Lys Pro Leu Ser Leu Asn  
                   355                                  360                                  365

gga aca act cag acg ttt ctc gaa gac gga gac caa gtc acc ttc tca 1152  
 Gly Thr Thr Gln Thr Phe Leu Glu Asp Gly Asp Gln Val Thr Phe Ser  
                   370                                  375                                  380

ggt gta tgc aag gga gat ggt tac aat gtt ggg ttt gga aca tgc aca 1200  
 Gly Val Cys Lys Gly Asp Gly Tyr Asn Val Gly Phe Gly Thr Cys Thr  
                   385                                  390                                  395                                  400

ggg aaa att gtt cct tca ccg cct tga 1227  
 Gly Lys Ile Val Pro Ser Pro Pro  
                                   405

&lt;210&gt; 5

&lt;211&gt; 408

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 5

Met Ala Leu Leu Lys Ser Phe Ile Asp Val Gly Ser Asp Ser His Phe  
   1                                  5                                  10                                  15

Pro Ile Gln Asn Leu Pro Tyr Gly Val Phe Lys Pro Glu Ser Asn Ser  
                   20                                  25                                  30

Thr Pro Arg Pro Ala Val Ala Ile Gly Asp Leu Val Leu Asp Leu Ser  
                   35                                  40                                  45

Ala Ile Ser Glu Ala Gly Leu Phe Asp Gly Leu Ile Leu Lys Asp Ala  
                   50                                  55                                  60

Asp Cys Phe Leu Gln Pro Asn Leu Asn Lys Phe Leu Ala Met Gly Arg  
                   65                                  70                                  75                                  80

Pro Ala Trp Lys Glu Ala Arg Ser Thr Leu Gln Arg Ile Leu Ser Phe  
                   85                                  90                                  95

Leu Leu Phe Gly Phe Lys Val Leu Val Leu Val Cys Phe His Ala Ala  
                   100                                  105                                  110

Asn Glu Pro Ile Leu Arg Asp Asn Asp Val Leu Arg Arg Lys Ser Phe  
                   115                                  120                                  125



His Gln Met Ser Lys Val Glu Met Ile Val Pro Met Val Ile Gly Asp  
 130 135 140  
 Tyr Thr Asp Phe Phe Ala Ser Met His His Ala Lys Asn Cys Gly Leu  
 145 150 155 160  
 Met Phe Arg Gly Pro Glu Asn Ala Ile Asn Pro Asn Trp Phe Arg Leu  
 165 170 175  
 Pro Ile Ala Tyr His Gly Arg Ala Ser Ser Ile Val Ile Ser Gly Thr  
 180 185 190  
 Asp Ile Ile Arg Pro Arg Gly Gln Gly His Pro Gln Gly Asn Ser Glu  
 195 200 205  
 Pro Tyr Phe Gly Pro Ser Lys Lys Leu Asp Phe Glu Leu Glu Met Ala  
 210 215 220  
 Ala Val Val Gly Pro Gly Asn Glu Leu Gly Lys Pro Ile Asp Val Asn  
 225 230 235 240  
 Asn Ala Ala Asp His Ile Phe Gly Leu Leu Leu Met Asn Asp Trp Ser  
 245 250 255  
 Ala Arg Asp Ile Gln Ala Trp Glu Tyr Val Pro Leu Gly Pro Phe Leu  
 260 265 270  
 Gly Lys Ser Phe Gly Thr Thr Ile Ser Pro Trp Ile Val Thr Leu Asp  
 275 280 285  
 Ala Leu Glu Pro Phe Gly Cys Gln Ala Pro Lys Gln Asp Pro Pro Pro  
 290 295 300  
 Leu Pro Tyr Leu Ala Glu Lys Glu Ser Val Asn Tyr Asp Ile Ser Leu  
 305 310 315 320  
 Glu Leu Ala His His Thr Val Asn Gly Cys Asn Leu Arg Pro Gly Asp  
 325 330 335  
 Leu Leu Gly Thr Gly Thr Ile Ser Gly Pro Glu Pro Asp Ser Tyr Gly  
 340 345 350  
 Cys Leu Leu Glu Leu Thr Trp Asn Gly Gln Lys Pro Leu Ser Leu Asn  
 355 360 365  
 Gly Thr Thr Gln Thr Phe Leu Glu Asp Gly Asp Gln Val Thr Phe Ser  
 370 375 380  
 Gly Val Cys Lys Gly Asp Gly Tyr Asn Val Gly Phe Gly Thr Cys Thr  
 385 390 395 400  
 Gly Lys Ile Val Pro Ser Pro Pro  
 405

10

```

<210> 6
<211> 1721
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> gene
<222> (9)..(1713)
<223> gene for maleylacetoacetate isomerase (MAAI)

<220>
<221> misc_feature
<222> (1)..(8)
<223> restriction site linker

<220>
<221> misc_feature
<222> (1714)..(1721)
<223> restriction site linker

<400> 6
atgtcgacat gtcttatggt accgattttt atcaggcgaa gttgaagctc tactcttact 60
ggagaagctc atgtgctcat cgcgtccgta tcgccctcac tttaaaaggt accagccaat 120
gattttatct tttcttctgt agcaattctt tgatctgaat ttgggtcttg ttcgattttc 180
attagggctt gattatgaat atataccggt taatttgctc aaaggggatc aatccgattc 240
agggtgcgtg tttctagggt atattgaact ttatttgaag taacattgta aagataagaa 300
tggtaagtaa ctgagatttc ttatgttaga cttagaagtt tattcgtttt gggtctctag 360
atttcaagaa gatcaatcca atgggcactg taccagcgct tgttgatggt gatgttgtga 420
ttaatgactc tttcgcaata ataatggtca gtagtaacac atccatttag tttgtttggt 480
tttgttgatg aaaaggaaca ttctgtttatt cgtcttggtg tttttcaaat ggacagtacc 540
tggatgataa gtatccggag ccaccgctgt taccaagtga ctaccataaa cgggcggtaa 600
attaccaggt atcttcgac ctttgtcttc agatgatgat gtgttgccat catctgcaa 660
accatgtagt taagtccaaa tgtagtgaac attatcagct ttagattgct agtgtgatcg 720
ttgttcttat ttgtatatatt tcaggcgacg agtattgtca tgtctgggtat acagcctcat 780
caaaatatgg ctctttttgt gagaagatga gattaatgta atggattcta ctaatggagg 840
ttctataaca aagcaaacat agttacattt tgtcattttt tttaacagag gtatctcgag 900
gacaagataa atgctgagga gaaaactgct tggattacta atgctatcac aaaaggattc 960
acaggtatga tatctctaata ctacctatac gtaatcaaga accaagacat atgttcaaaa 1020
tgtgattttg ttgatattgt ggttgtagac gtttataacg acctgtctga taatgtctca 1080
tatgtccttc agctctcgag aaactggttg tgagttgcgc tggaaaatac gcgactgggtg 1140
atgaagttta ctgtgtatgt ctctaaatct ccctggataa tctctatggt actactctct 1200
tctttattac aatgaagcat tgttttgcag gctgatcttt tcctagcacc acagatccac 1260
gcagcattca acagattcca tattaacatg gtacttttcc tcagctaata tcttctcctg 1320
gtacctagat attgcattgt atatccccc aaattccatg gaatccttga tcagagtttt 1380
aaggtagcat gaaccaaag ttatctctgt ctcacacttt cacattcaca gagtaacata 1440
gacgtaatac tcagtttcat aacttttttt cctcgcatca cttggttttc atctctacaa 1500
ttttgttgta taggaacct tcccgaactc tgcaagggtt tacgagtcac acaacgaact 1560
gcctgcattt caaaatgcag tcccggagaa gcaaccagat actccttcca ccatctgatt 1620
ctgtgaaccg taagcttctc tcagttctag ctcaataaaa tctcttagga aacaacaaca 1680
acaccttgaa cttaaatgta tcatatgaac cagggatcca t 1721

```

## 11

```

<210> 7
<211> 1238
<212> DNA
<213> Escherichia coli

<220>
<221> gene
<222> (7)..(1232)
<223> tyrA gene coding for bifunctional chorismate
        mutase / prephenate dehydrogenase

<220>
<221> CDS
<222> (25)..(1143)

<220>
<221> misc_feature
<222> (1)..(6)
<223> restriction site linker

<220>
<221> misc_feature
<222> (1233)..(1238)
<223> restriction site linker

<400> 7
cccggtggc ttaagaggtt tatt atg gtt gct gaa ttg acc gca tta cgc      51
                Met Val Ala Glu Leu Thr Ala Leu Arg
                        1                      5

gat caa att gat gaa gtc gat aaa gcg ctg ctg aat tta tta gcg aag      99
Asp Gln Ile Asp Glu Val Asp Lys Ala Leu Leu Asn Leu Leu Ala Lys
  10                15                20                25

cgt ctg gaa ctg gtt gct gaa gtg ggc gag gtg aaa agc cgc ttt gga      147
Arg Leu Glu Leu Val Ala Glu Val Gly Glu Val Lys Ser Arg Phe Gly
                30                35                40

ctg cct att tat gtt ccg gag cgc gag gca tct atg ttg gcc tcg cgt      195
Leu Pro Ile Tyr Val Pro Glu Arg Glu Ala Ser Met Leu Ala Ser Arg
                45                50                55

cgt gca gag gcg gaa gct ctg ggt gta ccg cca gat ctg att gag gat      243
Arg Ala Glu Ala Glu Ala Leu Gly Val Pro Pro Asp Leu Ile Glu Asp
                60                65                70

gtt ttg cgt cgg gtg atg cgt gaa tct tac tcc agt gaa aac gac aaa      291
Val Leu Arg Arg Val Met Arg Glu Ser Tyr Ser Ser Glu Asn Asp Lys
                75                80                85

gga ttt aaa aca ctt tgt ccg tca ctg cgt ccg gtg gtt atc gtc ggc      339
Gly Phe Lys Thr Leu Cys Pro Ser Leu Arg Pro Val Val Ile Val Gly
  90                95                100                105

```

## 12

ggt ggc ggt cag atg gga cgc ctg ttc gag aag atg ctg acc ctc tcg	387
Gly Gly Gly Gln Met Gly Arg Leu Phe Glu Lys Met Leu Thr Leu Ser	
110 115 120	
ggt tat cag gtg cgg att ctg gag caa cat gac tgg gat cga gcg gct	435
Gly Tyr Gln Val Arg Ile Leu Glu Gln His Asp Trp Asp Arg Ala Ala	
125 130 135	
gat att gtt gcc gat gcc gga atg gtg att gtt agt gtg cca atc cac	483
Asp Ile Val Ala Asp Ala Gly Met Val Ile Val Ser Val Pro Ile His	
140 145 150	
gtt act gag caa gtt att ggc aaa tta ccg cct tta ccg aaa gat tgt	531
Val Thr Glu Gln Val Ile Gly Lys Leu Pro Pro Leu Pro Lys Asp Cys	
155 160 165	
att ctg gtc gat ctg gca tca gtg aaa aat ggg cca tta cag gcc atg	579
Ile Leu Val Asp Leu Ala Ser Val Lys Asn Gly Pro Leu Gln Ala Met	
170 175 180 185	
ctg gtg gcg cat gat ggt ccg gtg ctg ggg cta cac ccg atg ttc ggt	627
Leu Val Ala His Asp Gly Pro Val Leu Gly Leu His Pro Met Phe Gly	
190 195 200	
ccg gac agc ggt agc ctg gca aag caa gtt gtg gtc tgg tgt gat gga	675
Pro Asp Ser Gly Ser Leu Ala Lys Gln Val Val Val Trp Cys Asp Gly	
205 210 215	
cgt aaa ccg gaa gca tac caa tgg ttt ctg gag caa att cag gtc tgg	723
Arg Lys Pro Glu Ala Tyr Gln Trp Phe Leu Glu Gln Ile Gln Val Trp	
220 225 230	
ggc gct cgg ctg cat cgt att agc gcc gtc gag cac gat cag aat atg	771
Gly Ala Arg Leu His Arg Ile Ser Ala Val Glu His Asp Gln Asn Met	
235 240 245	
gcg ttt att cag gca ctg cgc cac ttt gct act ttt gct tac ggg ctg	819
Ala Phe Ile Gln Ala Leu Arg His Phe Ala Thr Phe Ala Tyr Gly Leu	
250 255 260 265	
cac ctg gca gaa gaa aat gtt cag ctt gag caa ctt ctg gcg ctc tct	867
His Leu Ala Glu Glu Asn Val Gln Leu Glu Gln Leu Leu Ala Leu Ser	
270 275 280	
tcg ccg att tac cgc ctt gag ctg gcg atg gtc ggg cga ctg ttt gct	915
Ser Pro Ile Tyr Arg Leu Glu Leu Ala Met Val Gly Arg Leu Phe Ala	
285 290 295	
cag gat ccg cag ctt tat gcc gac atc att atg tcg tca gag cgt aat	963
Gln Asp Pro Gln Leu Tyr Ala Asp Ile Ile Met Ser Ser Glu Arg Asn	
300 305 310	

## 13

ctg gcg tta atc aaa cgt tac tat aag cgt ttc ggc gag gcg att gag 1011  
 Leu Ala Leu Ile Lys Arg Tyr Tyr Lys Arg Phe Gly Glu Ala Ile Glu  
 315 320 325

ttg ctg gag cag ggc gat aag cag gcg ttt att gac agt ttc cgc aag 1059  
 Leu Leu Glu Gln Gly Asp Lys Gln Ala Phe Ile Asp Ser Phe Arg Lys  
 330 335 340 345

gtg gag cac tgg ttc ggc gat tac gca cag cgt ttt cag agt gaa agc 1107  
 Val Glu His Trp Phe Gly Asp Tyr Ala Gln Arg Phe Gln Ser Glu Ser  
 350 355 360

cgc gtg tta ttg cgt cag gcg aat gac aat cgc cag taataatcca 1153  
 Arg Val Leu Leu Arg Gln Ala Asn Asp Asn Arg Gln  
 365 370

gtgccggatg attcacatca tccggcacct tttcatcagg ttggatcaac aggcactacg 1213

ttctcacttg ggtaacagcg tcgac 1238

<210> 8

<211> 373

<212> PRT

<213> Escherichia coli

<400> 8

Met Val Ala Glu Leu Thr Ala Leu Arg Asp Gln Ile Asp Glu Val Asp  
 1 5 10 15

Lys Ala Leu Leu Asn Leu Leu Ala Lys Arg Leu Glu Leu Val Ala Glu  
 20 25 30

Val Gly Glu Val Lys Ser Arg Phe Gly Leu Pro Ile Tyr Val Pro Glu  
 35 40 45

Arg Glu Ala Ser Met Leu Ala Ser Arg Arg Ala Glu Ala Glu Ala Leu  
 50 55 60

Gly Val Pro Pro Asp Leu Ile Glu Asp Val Leu Arg Arg Val Met Arg  
 65 70 75 80

Glu Ser Tyr Ser Ser Glu Asn Asp Lys Gly Phe Lys Thr Leu Cys Pro  
 85 90 95

Ser Leu Arg Pro Val Val Ile Val Gly Gly Gly Gln Met Gly Arg  
 100 105 110

Leu Phe Glu Lys Met Leu Thr Leu Ser Gly Tyr Gln Val Arg Ile Leu  
 115 120 125

Glu Gln His Asp Trp Asp Arg Ala Ala Asp Ile Val Ala Asp Ala Gly  
 130 135 140

Met Val Ile Val Ser Val Pro Ile His Val Thr Glu Gln Val Ile Gly  
 145 150 155 160

## 14

Lys Leu Pro Pro Leu Pro Lys Asp Cys Ile Leu Val Asp Leu Ala Ser  
 165 170 175  
 Val Lys Asn Gly Pro Leu Gln Ala Met Leu Val Ala His Asp Gly Pro  
 180 185 190  
 Val Leu Gly Leu His Pro Met Phe Gly Pro Asp Ser Gly Ser Leu Ala  
 195 200 205  
 Lys Gln Val Val Val Trp Cys Asp Gly Arg Lys Pro Glu Ala Tyr Gln  
 210 215 220  
 Trp Phe Leu Glu Gln Ile Gln Val Trp Gly Ala Arg Leu His Arg Ile  
 225 230 235 240  
 Ser Ala Val Glu His Asp Gln Asn Met Ala Phe Ile Gln Ala Leu Arg  
 245 250 255  
 His Phe Ala Thr Phe Ala Tyr Gly Leu His Leu Ala Glu Glu Asn Val  
 260 265 270  
 Gln Leu Glu Gln Leu Leu Ala Leu Ser Ser Pro Ile Tyr Arg Leu Glu  
 275 280 285  
 Leu Ala Met Val Gly Arg Leu Phe Ala Gln Asp Pro Gln Leu Tyr Ala  
 290 295 300  
 Asp Ile Ile Met Ser Ser Glu Arg Asn Leu Ala Leu Ile Lys Arg Tyr  
 305 310 315 320  
 Tyr Lys Arg Phe Gly Glu Ala Ile Glu Leu Leu Glu Gln Gly Asp Lys  
 325 330 335  
 Gln Ala Phe Ile Asp Ser Phe Arg Lys Val Glu His Trp Phe Gly Asp  
 340 345 350  
 Tyr Ala Gln Arg Phe Gln Ser Glu Ser Arg Val Leu Leu Arg Gln Ala  
 355 360 365  
 Asn Asp Asn Arg Gln  
 370

&lt;210&gt; 9

&lt;211&gt; 2953

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; gene

&lt;222&gt; (1)..(2953)

&lt;223&gt; gene for fumarylacetoacetate hydrolase (FAAH)

&lt;400&gt; 9

atggcgttgc tgaagtcttt catcgatgtt ggctcagact cgcacttccc tatccagaat 60  
 ctcccttatg gtgtcttcaa accggaatcg aactcaactc ctgctcctgc cgctcgctatc 120  
 ggcgatttgg ttctggacct ctccgctatc tctgaagctg ggcttttccg tgggtctgac 180  
 cttaaggacg cagattgctt tcttcagggt cgtttttccg attcctataa actcggatta 240  
 ctatgtagta gtaccctggg aatgtttccg taaatgattt cgaatttgct atttgaacct 300  
 gatctctgaa gtttgctcca tggtttattg gatagatcaa tcccgtttag ctcgaaaaaa 360  
 atccattgtt ctactcaatt gctcgttgct tcgattcatt atctgttaca gtttgagttt 420  
 tctgttcacg attttgaact ttgcaacta tgattattgc tttatgatct gacggtatag 480  
 tgtattgctt acacttagtg atgaggaaaa tgaggttgtg tttattttct ggtgtgtttc 540  
 ttttgatgtt aatattgttt agtttctgtg ctctgtttgc agcctaattt gaataagttc 600  
 ttggccatgg gacggcctgc gtggaaggaa gcgcgttcta cgctgcaaag aatcttgtca 660  
 tgtatgctct gtttgatcct attgatttat ttggattttt atggagtttt gttatttggc 720  
 ttcaaggttt tgggttttgg atgttttcat gcagctaata aacctatctt gcgagacaat 780  
 gatgttttga ggagaaaatc attccatcag atggttagta gtgtgaaatt gttttttgct 840  
 taaactaggg aaattgtttg tatatctgtt acttacgttt attgctgttt gatgcaaatt 900  
 tgcagagtaa agtggaaatg attgttcta tgggtgattg ggactataca gacttctttg 960  
 catctatgca tcacgcgaag aactgaggac ttatgttccg tgggcctgag aatgcgataa 1020  
 acccaaattg gtgcgtttat gttacttttg agctgagagt ttcttcatga aatgggtcaag 1080  
 tcgaaaggat gactctgtat taacatgaca ttaccataat tttcaggttt cgtcttccca 1140  
 ttgcatatca tggacgggca tcatctattg tcatctctgg gactgacatt attcgaccaa 1200  
 ggtaggaaa ttgtgtatta ttatctgggt tttggtgggc tgagaatggt tgttaagaat 1260  
 aattcacatg tcatatttga agtcatgcat catgcaagggt tttatgcttt gacaagaaat 1320  
 atagtttttt ataagatatt attacattga aaccaatatt ggcggtatggt aaaatttcat 1380  
 gcagacaaat taataatgaa atgctaattc cagttttatc tttgcttgtt ttgctttctt 1440  
 ccagagggtca gggccatcca caaggaaact ctgaaccata ttttgacact tcgaagaaac 1500  
 ttgattttga gcttgagatg gtaagcatct gatgcctcag ttatgtggat ttgttttaca 1560  
 atgattcgggt tgatgctttt tgggtgctagt taagaataac ggcatgaca aacctctctt 1620  
 ttatcacatg atattcaggc tgctgtgggt ggtccaggaa atgaattggg aaagcctatt 1680  
 gacgtgaata atgcagccga tcatataattt ggtctattac tgatgaatga ctggagtgg 1740  
 actcacttaa ctatagtttt cggtgagtca tctttaacct gaccgggcat gaccggtttt 1800  
 tttaaatgtt tgttgttata gctagggata ttcaggcgtg ggagtatgta cctcttgggtc 1860  
 ctttctggg gaagagtttt ggtgagatat ttggcttcaa tactttgatt tcatttctctc 1920  
 tagttgaagt atatgggcaa agaacttcgg tgaatgttgt cttgttgtgt tgtagggact 1980  
 actatatccc cttggattgt taccttggat gcgcttgagc cttttgggtg tcaagctccc 2040  
 aagcagggtg gtacttaggc atcacattct ttttgtgtca cgcaatcact gattctctca 2100  
 tgatctaact tgttcttggg gcaggatcca cctccattgc catatttggc tgagaaagag 2160  
 tctgtaaatt acgatatctc cttggaggta gcattcgata ttggagtttc actttttggc 2220  
 tttttgctat caactataac agcttatggg ggactgaact gaaataaaca tcatgttttt 2280  
 acctcttata ggttcaactt aaaccttctg gcagagatga ttcttgtgta ataacaaaga 2340  
 gcaacttcca aaacttgtga gttcctctat aatctctac ccaattctc catataatta 2400  
 aacagtttgg ttcaaaactt tttaaactta ttgtgacaga tattggacca taacgcagca 2460  
 gctagcacac cataccgtta acggttgcaa tttgaggcct ggtgatctcc ttggcacagg 2520  
 aaccataagc ggaccggtaa actcttttctg aaccagttct ctgctctact atatcacgtg 2580  
 atgactacac aataactcgc aaaatctttg tttcttgggt ctaaacgcag gagccagatt 2640  
 catatgggtg cctacttgag ttgacatgga atggacagaa acctctatca ctcaatggaa 2700  
 caactcagac gtttctcgaa gacggagacc aagtcacctt ctgaggtgta tgcaagggtat 2760  
 cagctgatta acacggtttc tgctttagtt taatttgcct tataccccaa caactccaaa 2820  
 tgaatttcgt tgcattgacat ttcggttaac gcttattaat caaattacgt ctatgattaa 2880  
 accgtttagt ggagatgggt acaatgttgg gtttggaaaca tgcacaggga aaattgtttc 2940

16

ttcaccgcct tga

2953

&lt;210&gt; 10

&lt;211&gt; 1534

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (28)..(1227)

&lt;223&gt; cDNA coding for hydroxyphenylpyruvate dioxygenase

&lt;400&gt; 10

cgagtttttag cagagttgggt gaaatca atg ggc cac caa aac gcc gcc gtt tca 54

Met Gly His Gln Asn Ala Ala Val Ser

1

5

gag aat caa aac cat gat gac ggc gct gcg tcg tcg ccg gga ttc aag 102

Glu Asn Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys

10

15

20

25

ctc gtc gga ttt tcc aag ttc gta aga aag aat cca aag tct gat aaa 150

Leu Val Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys

30

35

40

ttc aag gtt aag cgc ttc cat cac atc gag ttc tgg tgc ggc gac gca 198

Phe Lys Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala

45

50

55

acc aac gtc gct cgt cgc ttc tcc tgg ggt ctg ggg atg aga ttc tcc 246

Thr Asn Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser

60

65

70

gcc aaa tcc gat ctt tcc acc gga aac atg gtt cac gcc tct tac cta 294

Ala Lys Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu

75

80

85

ctc acc tcc ggt gac ctc cga ttc ctt ttc act gct cct tac tct ccg 342

Leu Thr Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro

90

95

100

105

tct ctc tcc gcc gga gag att aaa ccg aca acc aca gct tct atc cca 390

Ser Leu Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro

110

115

120

agt ttc gat cac ggc tct tgt cgt tcc ttc ttc tct tca cat ggt ctc 438

Ser Phe Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu

125

130

135

ggg gtt aga gcc gtt gcg att gaa gta gaa gac gca gag tca gct ttc 486

Gly Val Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe

140

145

150



## 17

tcc atc agt gta gct aat ggc gct att cct tcg tcg cct cct atc gtc	534
Ser Ile Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val	
155 160 165	
ctc aat gaa gca gtt acg atc gct gag gtt aaa cta tac ggc gat gtt	582
Leu Asn Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val	
170 175 180 185	
gtt ctc cga tat gtt agt tac aaa gca gaa gat acc gaa aaa tcc gaa	630
Val Leu Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu	
190 195 200	
ttc ttg cca ggg ttc gag cgt gta gag gat gcg tcg tcg ttc cca ttg	678
Phe Leu Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu	
205 210 215	
gat tat ggt atc cgg cgg ctt gac cac gcc gtg gga aac gtt cct gag	726
Asp Tyr Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu	
220 225 230	
ctt ggt ccg gct tta act tat gta gcg ggg ttc act ggt ttt cac caa	774
Leu Gly Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln	
235 240 245	
ttc gca gag ttc aca gca gac gac gtt gga acc gcc gag agc ggt tta	822
Phe Ala Glu Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu	
250 255 260 265	
aat tca gcg gtc ctg gct agc aat gat gaa atg gtt ctt cta ccg att	870
Asn Ser Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile	
270 275 280	
aac gag cca gtg cac gga aca aag agg aag agt cag att cag acg tat	918
Asn Glu Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr	
285 290 295	
ttg gaa cat aac gaa ggc gca ggg cta caa cat ctg gct ctg atg agt	966
Leu Glu His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser	
300 305 310	
gaa gac ata ttc agg acc ctg aga gag atg agg aag agg agc agt att	1014
Glu Asp Ile Phe Arg Thr Leu Arg Glu Met Arg Lys Arg Ser Ser Ile	
315 320 325	
gga gga ttc gac ttc atg cct tct cct ccg cct act tac tac cag aat	1062
Gly Gly Phe Asp Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr Gln Asn	
330 335 340 345	
ctc aag aaa ccg gtc ggc gac gtg ctc agc gat gat cag atc aag gag	1110
Leu Lys Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu	
350 355 360	

## 18

tgt gag gaa tta ggg att ctt gta gac aga gat gat caa ggg acg ttg 1158  
 Cys Glu Glu Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly Thr Leu  
                   365                                  370                                  375

ctt caa atc ttc aca aaa cca cta ggt gac agg tac agt tca ttt aat 1206  
 Leu Gln Ile Phe Thr Lys Pro Leu Gly Asp Arg Tyr Ser Ser Phe Asn  
                   380                                  385                                  390

caa aca cat gtt aca gtt ccc taacaatcca ttgatgata aacatgttac 1257  
 Gln Thr His Val Thr Val Pro  
                   395                                  400

agtttactaa gcaatctctt gtttatgatt gtgttaatag gccgacgata tttatagaga 1317  
 taatccagag agtaggatgc atgatgaaag atgaggaagg gaaggcttac cagagtggag 1377  
 gatgtggtgg tctctgagct cttcaagtcc attgaagaat acgaaaagac tcttgaagcc 1437  
 aaacagttag tgggatgaac aagaagaaga accaactaaa ggattgtgta attaatgtaa 1497  
 aactgtttta tcttatcaaa acaatgttat acaacat 1534

<210> 11

<211> 400

<212> PRT

<213> Arabidopsis thaliana

<400> 11

Met Gly His Gln Asn Ala Ala Val Ser Glu Asn Gln Asn His Asp Asp  
           1                                  5                                  10                                  15

Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val Gly Phe Ser Lys Phe  
                   20                                  25                                  30

Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys Val Lys Arg Phe His  
                   35                                  40                                  45

His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn Val Ala Arg Arg Phe  
           50                                  55                                  60

Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys Ser Asp Leu Ser Thr  
           65                                  70                                  75                                  80

Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr Ser Gly Asp Leu Arg  
                   85                                  90                                  95

Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu Ser Ala Gly Glu Ile  
                   100                                  105                                  110

Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe Asp His Gly Ser Cys  
           115                                  120                                  125

Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val Arg Ala Val Ala Ile  
           130                                  135                                  140

## 19

Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile Ser Val Ala Asn Gly  
 145 150 155 160  
 Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn Glu Ala Val Thr Ile  
 165 170 175  
 Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu Arg Tyr Val Ser Tyr  
 180 185 190  
 Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu Pro Gly Phe Glu Arg  
 195 200 205  
 Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr Gly Ile Arg Arg Leu  
 210 215 220  
 Asp His Ala Val Gly Asn Val Pro Glu Leu Gly Pro Ala Leu Thr Tyr  
 225 230 235 240  
 Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala Glu Phe Thr Ala Asp  
 245 250 255  
 Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser Ala Val Leu Ala Ser  
 260 265 270  
 Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu Pro Val His Gly Thr  
 275 280 285  
 Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu His Asn Glu Gly Ala  
 290 295 300  
 Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp Ile Phe Arg Thr Leu  
 305 310 315 320  
 Arg Glu Met Arg Lys Arg Ser Ser Ile Gly Gly Phe Asp Phe Met Pro  
 325 330 335  
 Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys Lys Arg Val Gly Asp  
 340 345 350  
 Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu Glu Leu Gly Ile Leu  
 355 360 365  
 Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln Ile Phe Thr Lys Pro  
 370 375 380  
 Leu Gly Asp Arg Tyr Ser Ser Phe Asn Gln Thr His Val Thr Val Pro  
 385 390 395 400

&lt;210&gt; 12

&lt;211&gt; 575

&lt;212&gt; DNA

&lt;213&gt; Brassica napus

20

<220>  
 <221> misc\_feature  
 <222> (1)..(6)  
 <223> restriction site

<220>  
 <221> misc\_feature  
 <222> (570)..(575)  
 <223> restriction site

<220>  
 <221> misc\_feature  
 <222> (7)..(569)  
 <223> fragment of cDNA coding for  
 homogentisat-1,2-dioxygenase

<400> 12  
 gtcgacgggc cgatgggggc gaagggtctt gctgcaccaa gagattttct tgcaccaacg 60  
 gcatggtttg aggaagggtc acggcctgac tacactattg ttcagaagtt tggcggtgaa 120  
 ctctttactg ctaaacaaga tttctctccg ttcaatgtgg ttgcctggca tggcaattac 180  
 gtgccttata agtatgacct gcacaagttc tgtccatata acaactgtcct tgtagaccat 240  
 ggagatccat ctgtaaatac agttctgaca gcaccaacgg ataaacctgg tgtggccttg 300  
 cttgattttg tcatattccc tctcgttgg ttggttgctg agcatacctt tcgacctcct 360  
 tactaccatc gtaactgcat gagtgaattt atgggcctaa tctatggtgc ttacgaggcc 420  
 aaagctgatg gattttctacc tgggtggcgca agtcttcaca gttgtatgac acctcatggt 480  
 ccagatacaa ccacatacga ggcgacgatt gctcgtgtaa atgcaatggc tccttataag 540  
 ctcacaggca ccatggcctt catgtttgag gtacc 575

<210> 13  
 <211> 932  
 <212> DNA  
 <213> Synechocystis PCC6803

<220>  
 <221> CDS  
 <222> (4)..(927)  
 <223> cDNA coding for homogentisat phytyltransferase

<400> 13  
 gcc atg gca act atc caa gct ttt tgg cgc ttc tcc cgc ccc cat acc 48  
 Met Ala Thr Ile Gln Ala Phe Trp Arg Phe Ser Arg Pro His Thr  
 1 5 10 15  
 atc att ggt aca act ctg agc gtc tgg gct gtg tat ctg tta act att 96  
 Ile Ile Gly Thr Thr Leu Ser Val Trp Ala Val Tyr Leu Leu Thr Ile  
 20 25 30  
 ctc ggg gat gga aac tca gtt aac tcc cct gct tcc ctg gat tta gtg 144  
 Leu Gly Asp Gly Asn Ser Val Asn Ser Pro Ala Ser Leu Asp Leu Val  
 35 40 45

## 21

ttc ggc gct tgg ctg gcc tgc ctg ttg ggt aat gtg tac att gtc ggc	192
Phe Gly Ala Trp Leu Ala Cys Leu Leu Gly Asn Val Tyr Ile Val Gly	
50 55 60	
ctc aac caa ttg tgg gat gtg gac att gac cgc atc aat aag ccg aat	240
Leu Asn Gln Leu Trp Asp Val Asp Ile Asp Arg Ile Asn Lys Pro Asn	
65 70 75	
ttg ccc cta gct aac gga gat ttt tct atc gcc cag ggc cgt tgg att	288
Leu Pro Leu Ala Asn Gly Asp Phe Ser Ile Ala Gln Gly Arg Trp Ile	
80 85 90 95	
gtg gga ctt tgt ggc gtt gct tcc ttg gcg atc gcc tgg gga tta ggg	336
Val Gly Leu Cys Gly Val Ala Ser Leu Ala Ile Ala Trp Gly Leu Gly	
100 105 110	
cta tgg ctg ggg cta acg gtg ggc att agt ttg att att ggc acg gcc	384
Leu Trp Leu Gly Leu Thr Val Gly Ile Ser Leu Ile Ile Gly Thr Ala	
115 120 125	
tat tcg gtg ccg cca gtg agg tta aag cgc ttt tcc ctg ctg gcg gcc	432
Tyr Ser Val Pro Pro Val Arg Leu Lys Arg Phe Ser Leu Leu Ala Ala	
130 135 140	
ctg tgt att ctg acg gtg cgg gga att gtg gtt aac ttg ggc tta ttt	480
Leu Cys Ile Leu Thr Val Arg Gly Ile Val Val Asn Leu Gly Leu Phe	
145 150 155	
tta ttt ttt aga att ggt tta ggt tat ccc ccc act tta ata acc ccc	528
Leu Phe Phe Arg Ile Gly Leu Gly Tyr Pro Pro Thr Leu Ile Thr Pro	
160 165 170 175	
atc tgg gtt ttg act tta ttt atc tta gtt ttc acc gtg gcg atc gcc	576
Ile Trp Val Leu Thr Leu Phe Ile Leu Val Phe Thr Val Ala Ile Ala	
180 185 190	
att ttt aaa gat gtg cca gat atg gaa ggc gat cgg caa ttt aag att	624
Ile Phe Lys Asp Val Pro Asp Met Glu Gly Asp Arg Gln Phe Lys Ile	
195 200 205	
caa act tta act ttg caa atc ggc aaa caa aac gtt ttt cgg gga acc	672
Gln Thr Leu Thr Leu Gln Ile Gly Lys Gln Asn Val Phe Arg Gly Thr	
210 215 220	
tta att tta ctc act ggt tgt tat tta gcc atg gca atc tgg ggc tta	720
Leu Ile Leu Leu Thr Gly Cys Tyr Leu Ala Met Ala Ile Trp Gly Leu	
225 230 235	
tgg gcg gct atg cct tta aat act gct ttc ttg att gtt tcc cat ttg	768
Trp Ala Ala Met Pro Leu Asn Thr Ala Phe Leu Ile Val Ser His Leu	
240 245 250 255	

## 22

tgc tta tta gcc tta ctc tgg tgg cgg agt cga gat gta cac tta gaa 816  
 Cys Leu Leu Ala Leu Leu Trp Trp Arg Ser Arg Asp Val His Leu Glu  
                   260                  265                  270

agc aaa acc gaa att gct agt ttt tat cag ttt att tgg aag cta ttt 864  
 Ser Lys Thr Glu Ile Ala Ser Phe Tyr Gln Phe Ile Trp Lys Leu Phe  
                   275                  280                  285

ttc tta gag tac ttg ctg tat ccc ttg gct ctg tgg tta cct aat ttt 912  
 Phe Leu Glu Tyr Leu Leu Tyr Pro Leu Ala Leu Trp Leu Pro Asn Phe  
                   290                  295                  300

tct aat act att ttt taggg 932  
 Ser Asn Thr Ile Phe  
                   305

<210> 14

<211> 308

<212> PRT

<213> Synechocystis PCC6803

<400> 14

Met Ala Thr Ile Gln Ala Phe Trp Arg Phe Ser Arg Pro His Thr Ile  
   1                  5                  10                  15

Ile Gly Thr Thr Leu Ser Val Trp Ala Val Tyr Leu Leu Thr Ile Leu  
                   20                  25                  30

Gly Asp Gly Asn Ser Val Asn Ser Pro Ala Ser Leu Asp Leu Val Phe  
                   35                  40                  45

Gly Ala Trp Leu Ala Cys Leu Leu Gly Asn Val Tyr Ile Val Gly Leu  
                   50                  55                  60

Asn Gln Leu Trp Asp Val Asp Ile Asp Arg Ile Asn Lys Pro Asn Leu  
                   65                  70                  75                  80

Pro Leu Ala Asn Gly Asp Phe Ser Ile Ala Gln Gly Arg Trp Ile Val  
                   85                  90                  95

Gly Leu Cys Gly Val Ala Ser Leu Ala Ile Ala Trp Gly Leu Gly Leu  
                   100                  105                  110

Trp Leu Gly Leu Thr Val Gly Ile Ser Leu Ile Ile Gly Thr Ala Tyr  
                   115                  120                  125

Ser Val Pro Pro Val Arg Leu Lys Arg Phe Ser Leu Leu Ala Ala Leu  
                   130                  135                  140

Cys Ile Leu Thr Val Arg Gly Ile Val Val Asn Leu Gly Leu Phe Leu  
                   145                  150                  155                  160

Phe Phe Arg Ile Gly Leu Gly Tyr Pro Pro Thr Leu Ile Thr Pro Ile  
                   165                  170                  175

## 23

Trp Val Leu Thr Leu Phe Ile Leu Val Phe Thr Val Ala Ile Ala Ile  
 180 185 190

Phe Lys Asp Val Pro Asp Met Glu Gly Asp Arg Gln Phe Lys Ile Gln  
 195 200 205

Thr Leu Thr Leu Gln Ile Gly Lys Gln Asn Val Phe Arg Gly Thr Leu  
 210 215 220

Ile Leu Leu Thr Gly Cys Tyr Leu Ala Met Ala Ile Trp Gly Leu Trp  
 225 230 235 240

Ala Ala Met Pro Leu Asn Thr Ala Phe Leu Ile Val Ser His Leu Cys  
 245 250 255

Leu Leu Ala Leu Leu Trp Trp Arg Ser Arg Asp Val His Leu Glu Ser  
 260 265 270

Lys Thr Glu Ile Ala Ser Phe Tyr Gln Phe Ile Trp Lys Leu Phe Phe  
 275 280 285

Leu Glu Tyr Leu Leu Tyr Pro Leu Ala Leu Trp Leu Pro Asn Phe Ser  
 290 295 300

Asn Thr Ile Phe  
 305

<210> 15

<211> 1159

<212> DNA

<213> Knnstliche Sequenz

<220>

<221> CDS

<222> (8)..(1150)

<220>

<221> misc\_feature

<222> (1)..(6)

<223> restriction site

<220>

<221> misc\_feature

<222> (1154)..(1159)

<223> restriction site

<220>

<223> Beschreibung der knnstlichen Sequenz: codon usage  
 optimized cDNA coding for hydroxyphenylpyruvate  
 dioxygenase from Streptomyces avermitilis

<400> 15

gtcgact atg act caa act act cat cat act cca gat act gct aga caa 49  
 Met Thr Gln Thr Thr His His Thr Pro Asp Thr Ala Arg Gln  
 1 5 10

24

gct gat cct ttt cca gtt aag gga atg gat gct gtt gtt ttc gct gtt	97
Ala Asp Pro Phe Pro Val Lys Gly Met Asp Ala Val Val Phe Ala Val	
15 20 25 30	
gga aac gct aag caa gct gct cat tac tac tct act gct ttc gga atg	145
Gly Asn Ala Lys Gln Ala Ala His Tyr Tyr Ser Thr Ala Phe Gly Met	
35 40 45	
caa ctt gtt gct tac tct gga cca gaa aac gga tct aga gaa act gct	193
Gln Leu Val Ala Tyr Ser Gly Pro Glu Asn Gly Ser Arg Glu Thr Ala	
50 55 60	
tct tac gtt ctt act aac gga tct gct aga ttc gtt ctt act tct gtt	241
Ser Tyr Val Leu Thr Asn Gly Ser Ala Arg Phe Val Leu Thr Ser Val	
65 70 75	
att aag cca gct acc cca tgg gga cat ttc ctt gct gat cac gtt gct	289
Ile Lys Pro Ala Thr Pro Trp Gly His Phe Leu Ala Asp His Val Ala	
80 85 90	
gaa cac gga gat gga gtt gtt gat ctt gct att gaa gtt cca gat gct	337
Glu His Gly Asp Gly Val Val Asp Leu Ala Ile Glu Val Pro Asp Ala	
95 100 105 110	
aga gct gct cat gct tac gct att gaa cat gga gct aga tct gtt gct	385
Arg Ala Ala His Ala Tyr Ala Ile Glu His Gly Ala Arg Ser Val Ala	
115 120 125	
gaa cca tac gaa ctt aag gat gaa cat gga act gtt gtt ctt gct gct	433
Glu Pro Tyr Glu Leu Lys Asp Glu His Gly Thr Val Val Leu Ala Ala	
130 135 140	
att gct act tac gga aag act aga cat act ctt gtt gat aga act gga	481
Ile Ala Thr Tyr Gly Lys Thr Arg His Thr Leu Val Asp Arg Thr Gly	
145 150 155	
tac gat gga cca tac ctt cca gga tac gtt gct gct gct cca att gtt	529
Tyr Asp Gly Pro Tyr Leu Pro Gly Tyr Val Ala Ala Ala Pro Ile Val	
160 165 170	
gaa cca cca gct cat aga acc ttc caa gct att gac cat tgt gtt ggt	577
Glu Pro Pro Ala His Arg Thr Phe Gln Ala Ile Asp His Cys Val Gly	
175 180 185 190	
aac gtt gaa ctc gga aga atg aac gaa tgg gtt gga ttc tac aac aag	625
Asn Val Glu Leu Gly Arg Met Asn Glu Trp Val Gly Phe Tyr Asn Lys	
195 200 205	
gtt atg gga ttc act aac atg aag gaa ttc gtt gga gat gat att gct	673
Val Met Gly Phe Thr Asn Met Lys Glu Phe Val Gly Asp Asp Ile Ala	
210 215 220	



## 25

```

act gag tac tct gct ctt atg tct aag gtt gtt gct gat gga act ctt 721
Thr Glu Tyr Ser Ala Leu Met Ser Lys Val Val Ala Asp Gly Thr Leu
      225              230              235

aag gtt aaa ttc cca att aat gaa cca gct ctt gct aag aag aag tct 769
Lys Val Lys Phe Pro Ile Asn Glu Pro Ala Leu Ala Lys Lys Lys Ser
      240              245              250

cag att gat gaa tac ctt gag ttc tac gga gga gct gga gtt caa cat 817
Gln Ile Asp Glu Tyr Leu Glu Phe Tyr Gly Gly Ala Gly Val Gln His
      255              260              265              270

att gct ctt aac act gga gat atc gtg gaa act gtt aga act atg aga 865
Ile Ala Leu Asn Thr Gly Asp Ile Val Glu Thr Val Arg Thr Met Arg
      275              280              285

gct gca gga gtt caa ttc ctt gat act cca gat tct tac tac gat act 913
Ala Ala Gly Val Gln Phe Leu Asp Thr Pro Asp Ser Tyr Tyr Asp Thr
      290              295              300

ctt ggt gaa tgg gtt gga gat act aga gtt cca gtt gat act ctt aga 961
Leu Gly Glu Trp Val Gly Asp Thr Arg Val Pro Val Asp Thr Leu Arg
      305              310              315

gaa ctt aag att ctt gct gat aga gat gaa gat gga tac ctt ctt caa 1009
Glu Leu Lys Ile Leu Ala Asp Arg Asp Glu Asp Gly Tyr Leu Leu Gln
      320              325              330

atc ttc act aag cca gtt caa gat aga cca act gtg ttc ttc gaa atc 1057
Ile Phe Thr Lys Pro Val Gln Asp Arg Pro Thr Val Phe Phe Glu Ile
      335              340              345              350

att gaa aga cat gga tct atg gga ttc gga aag ggt aac ttc aag gct 1105
Ile Glu Arg His Gly Ser Met Gly Phe Gly Lys Gly Asn Phe Lys Ala
      355              360              365

ctt ttc gaa gct att gaa aga gaa caa gag aag aga gga aac ctt 1150
Leu Phe Glu Ala Ile Glu Arg Glu Gln Glu Lys Arg Gly Asn Leu
      370              375              380

taggtcgac 1159

<210> 16
<211> 381
<212> PRT
<213> Kunstliche Sequenz
<223> Beschreibung der knnstlichen Sequenz: codon usage
      optimized cDNA coding for hydroxyphenylpyruvate
      dioxygenase from Streptomyces avermitilis

<400> 16
Met Thr Gln Thr Thr His His Thr Pro Asp Thr Ala Arg Gln Ala Asp
  1              5              10              15

```

26

Pro Phe Pro Val Lys Gly Met Asp Ala Val Val Phe Ala Val Gly Asn  
                   20                  25                  30  
 Ala Lys Gln Ala Ala His Tyr Tyr Ser Thr Ala Phe Gly Met Gln Leu  
           35                  40                  45  
 Val Ala Tyr Ser Gly Pro Glu Asn Gly Ser Arg Glu Thr Ala Ser Tyr  
           50                  55                  60  
 Val Leu Thr Asn Gly Ser Ala Arg Phe Val Leu Thr Ser Val Ile Lys  
           65                  70                  75                  80  
 Pro Ala Thr Pro Trp Gly His Phe Leu Ala Asp His Val Ala Glu His  
                   85                  90                  95  
 Gly Asp Gly Val Val Asp Leu Ala Ile Glu Val Pro Asp Ala Arg Ala  
                   100                  105                  110  
 Ala His Ala Tyr Ala Ile Glu His Gly Ala Arg Ser Val Ala Glu Pro  
           115                  120                  125  
 Tyr Glu Leu Lys Asp Glu His Gly Thr Val Val Leu Ala Ala Ile Ala  
           130                  135                  140  
 Thr Tyr Gly Lys Thr Arg His Thr Leu Val Asp Arg Thr Gly Tyr Asp  
           145                  150                  155                  160  
 Gly Pro Tyr Leu Pro Gly Tyr Val Ala Ala Ala Pro Ile Val Glu Pro  
                   165                  170                  175  
 Pro Ala His Arg Thr Phe Gln Ala Ile Asp His Cys Val Gly Asn Val  
                   180                  185                  190  
 Glu Leu Gly Arg Met Asn Glu Trp Val Gly Phe Tyr Asn Lys Val Met  
           195                  200                  205  
 Gly Phe Thr Asn Met Lys Glu Phe Val Gly Asp Asp Ile Ala Thr Glu  
           210                  215                  220  
 Tyr Ser Ala Leu Met Ser Lys Val Val Ala Asp Gly Thr Leu Lys Val  
           225                  230                  235                  240  
 Lys Phe Pro Ile Asn Glu Pro Ala Leu Ala Lys Lys Lys Ser Gln Ile  
                   245                  250                  255  
 Asp Glu Tyr Leu Glu Phe Tyr Gly Gly Ala Gly Val Gln His Ile Ala  
           260                  265                  270  
 Leu Asn Thr Gly Asp Ile Val Glu Thr Val Arg Thr Met Arg Ala Ala  
           275                  280                  285  
 Gly Val Gln Phe Leu Asp Thr Pro Asp Ser Tyr Tyr Asp Thr Leu Gly  
           290                  295                  300  
 Glu Trp Val Gly Asp Thr Arg Val Pro Val Asp Thr Leu Arg Glu Leu  
           305                  310                  315                  320

Lys	Ile	Leu	Ala	Asp	Arg	Asp	Glu	Asp	Gly	Tyr	Leu	Leu	Gln	Ile	Phe
				325					330					335	
Thr	Lys	Pro	Val	Gln	Asp	Arg	Pro	Thr	Val	Phe	Phe	Glu	Ile	Ile	Glu
			340					345					350		
Arg	His	Gly	Ser	Met	Gly	Phe	Gly	Lys	Gly	Asn	Phe	Lys	Ala	Leu	Phe
		355					360					365			
Glu	Ala	Ile	Glu	Arg	Glu	Gln	Glu	Lys	Arg	Gly	Asn	Leu			
	370					375					380				

aat tac cag gcg acg agt att gtc atg tct ggt ata cag cct cat caa 390  
Asn Tyr Gln Ala Thr Ser Ile Val Met Ser Gly Ile Gln Pro His Gln  
105 110 115

## 28

aat atg gct ctt ttt agg tat ctc gag gac aag ata aat gct gag gag 438  
 Asn Met Ala Leu Phe Arg Tyr Leu Glu Asp Lys Ile Asn Ala Glu Glu  
 120 125 130

aaa act gct tgg att act aat gct atc aca aaa gga ttc aca gct ctc 486  
 Lys Thr Ala Trp Ile Thr Asn Ala Ile Thr Lys Gly Phe Thr Ala Leu  
 135 140 145 150

gag aaa ctg ttg gtg agt tgc gct gga aaa tac gcg act ggt gat gaa 534  
 Glu Lys Leu Leu Val Ser Cys Ala Gly Lys Tyr Ala Thr Gly Asp Glu  
 155 160 165

gtt tac ttg gct gat ctt ttc cta gca cca cag atc cac gca gca ttc 582  
 Val Tyr Leu Ala Asp Leu Phe Leu Ala Pro Gln Ile His Ala Ala Phe  
 170 175 180

aac aga ttc cat att aac atg gaa cca ttc ccg act ctt gca agg ttt 630  
 Asn Arg Phe His Ile Asn Met Glu Pro Phe Pro Thr Leu Ala Arg Phe  
 185 190 195

tac gag tca tac aac gaa ctg cct gca ttt caa aat gca gtc ccg gag 678  
 Tyr Glu Ser Tyr Asn Glu Leu Pro Ala Phe Gln Asn Ala Val Pro Glu  
 200 205 210

aag caa cca gat act cct tcc acc atc tgattctgtg aaccgtaagc 725  
 Lys Gln Pro Asp Thr Pro Ser Thr Ile  
 215 220

ttctctcagt ctcagctcaa taaaatctct taggaaacaa caacaacacc ttgaacttaa 785

atgtatcata tgaaccagtt tacaataat 815

<210> 18

<211> 223

<212> PRT

<213> Arabidopsis thaliana

<400> 18

Met Ser Tyr Val Thr Asp Phe Tyr Gln Ala Lys Leu Lys Leu Tyr Ser  
 1 5 10 15

Tyr Trp Arg Ser Ser Cys Ala His Arg Val Arg Ile Ala Leu Thr Leu  
 20 25 30

Lys Gly Leu Asp Tyr Glu Tyr Ile Pro Val Asn Leu Leu Lys Gly Asp  
 35 40 45

Gln Ser Asp Ser Asp Phe Lys Lys Ile Asn Pro Met Gly Thr Val Pro  
 50 55 60

Ala Leu Val Asp Gly Asp Val Val Ile Asn Asp Ser Phe Ala Ile Ile  
 65 70 75 80

Met Tyr Leu Asp Asp Lys Tyr Pro Glu Pro Pro Leu Leu Pro Ser Asp  
 85 90 95

29

Tyr His Lys Arg Ala Val Asn Tyr Gln Ala Thr Ser Ile Val Met Ser  
 100 105 110  
 Gly Ile Gln Pro His Gln Asn Met Ala Leu Phe Arg Tyr Leu Glu Asp  
 115 120 125  
 Lys Ile Asn Ala Glu Glu Lys Thr Ala Trp Ile Thr Asn Ala Ile Thr  
 130 135 140  
 Lys Gly Phe Thr Ala Leu Glu Lys Leu Leu Val Ser Cys Ala Gly Lys  
 145 150 155 160  
 Tyr Ala Thr Gly Asp Glu Val Tyr Leu Ala Asp Leu Phe Leu Ala Pro  
 165 170 175  
 Gln Ile His Ala Ala Phe Asn Arg Phe His Ile Asn Met Glu Pro Phe  
 180 185 190  
 Pro Thr Leu Ala Arg Phe Tyr Glu Ser Tyr Asn Glu Leu Pro Ala Phe  
 195 200 205  
 Gln Asn Ala Val Pro Glu Lys Gln Pro Asp Thr Pro Ser Thr Ile  
 210 215 220

&lt;210&gt; 19

&lt;211&gt; 1350

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (63)..(1106)

&lt;223&gt; coding for gamma-tocopherol methyltransferase

&lt;400&gt; 19

ccacgcgtcc gcaaataatc cctgacttcg tcacgtttct ttgtatctcc aacgtccaat 60  
 aa atg aaa gca act cta gca gca ccc tct tct ctc aca agc ctc cct 107  
 Met Lys Ala Thr Leu Ala Ala Pro Ser Ser Leu Thr Ser Leu Pro  
 1 5 10 15  
 tat cga acc aac tct tct ttc ggc tca aag tca tcg ctt ctc ttt cgg 155  
 Tyr Arg Thr Asn Ser Ser Phe Gly Ser Lys Ser Ser Leu Leu Phe Arg  
 20 25 30  
 tct cca tcc tcc tcc tcc tca gtc tct atg acg aca acg cgt gga aac 203  
 Ser Pro Ser Ser Ser Ser Val Ser Met Thr Thr Thr Arg Gly Asn  
 35 40 45  
 gtg gct gtg gcg gct gct gct aca tcc act gag gcg cta aga aaa gga 251  
 Val Ala Val Ala Ala Ala Ala Thr Ser Thr Glu Ala Leu Arg Lys Gly  
 50 55 60

30

ata gcg gag ttc tac aat gaa act tcg ggt ttg tgg gaa gag att tgg	299
Ile Ala Glu Phe Tyr Asn Glu Thr Ser Gly Leu Trp Glu Glu Ile Trp	
65 70 75	
gga gat cat atg cat cat ggc ttt tat gac cct gat tct tct gtt caa	347
Gly Asp His Met His His Gly Phe Tyr Asp Pro Asp Ser Ser Val Gln	
80 85 90 95	
ctt tct gat tct ggt cac aag gaa gct cag atc cgt atg att gaa gag	395
Leu Ser Asp Ser Gly His Lys Glu Ala Gln Ile Arg Met Ile Glu Glu	
100 105 110	
tct ctc cgt ttc gcc ggt gtt act gat gaa gag gag gag aaa aag ata	443
Ser Leu Arg Phe Ala Gly Val Thr Asp Glu Glu Glu Glu Lys Lys Ile	
115 120 125	
aag aaa gta gtg gat gtt ggg tgt ggg att gga gga agc tca aga tat	491
Lys Lys Val Val Asp Val Gly Cys Gly Ile Gly Gly Ser Ser Arg Tyr	
130 135 140	
ctt gcc tct aaa ttt gga gct gaa tgc att ggc att act ctc agc cct	539
Leu Ala Ser Lys Phe Gly Ala Glu Cys Ile Gly Ile Thr Leu Ser Pro	
145 150 155	
gtt cag gcc aag aga gcc aat gat ctc gcg gct gct caa tca ctc tct	587
Val Gln Ala Lys Arg Ala Asn Asp Leu Ala Ala Ala Gln Ser Leu Ser	
160 165 170 175	
cat aag gct tcc ttc caa gtt gcg gat gcg ttg gat cag cca ttc gaa	635
His Lys Ala Ser Phe Gln Val Ala Asp Ala Leu Asp Gln Pro Phe Glu	
180 185 190	
gat gga aaa ttc gat cta gtg tgg tcg atg gag agt ggt gag cat atg	683
Asp Gly Lys Phe Asp Leu Val Trp Ser Met Glu Ser Gly Glu His Met	
195 200 205	
cct gac aag gcc aag ttt gta aaa gag ttg gta cgt gtg gcg gct cca	731
Pro Asp Lys Ala Lys Phe Val Lys Glu Leu Val Arg Val Ala Ala Pro	
210 215 220	
gga ggt agg ata ata ata gtg aca tgg tgc cat aga aat cta tct gcg	779
Gly Gly Arg Ile Ile Ile Val Thr Trp Cys His Arg Asn Leu Ser Ala	
225 230 235	
ggg gag gaa gct ttg cag ccg tgg gag caa aac atc ttg gac aaa atc	827
Gly Glu Glu Ala Leu Gln Pro Trp Glu Gln Asn Ile Leu Asp Lys Ile	
240 245 250 255	
tgt aag acg ttc tat ctc ccg gct tgg tgc tcc acc gat gat tat gtc	875
Cys Lys Thr Phe Tyr Leu Pro Ala Trp Cys Ser Thr Asp Asp Tyr Val	
260 265 270	

## 31

aac ttg ctt caa tcc cat tct ctc cag gat att aag tgt gcg gat tgg 923  
 Asn Leu Leu Gln Ser His Ser Leu Gln Asp Ile Lys Cys Ala Asp Trp  
 275 280 285

tca gag aac gta gct cct ttc tgg cct gcg gtt ata cgg act gca tta 971  
 Ser Glu Asn Val Ala Pro Phe Trp Pro Ala Val Ile Arg Thr Ala Leu  
 290 295 300

aca tgg aag ggc ctt gtg tct ctg ctt cgt agt ggt atg aaa agt att 1019  
 Thr Trp Lys Gly Leu Val Ser Leu Leu Arg Ser Gly Met Lys Ser Ile  
 305 310 315

aaa gga gca ttg aca atg cca ttg atg att gaa ggt tac aag aaa ggt 1067  
 Lys Gly Ala Leu Thr Met Pro Leu Met Ile Glu Gly Tyr Lys Lys Gly  
 320 325 330 335

gtc att aag ttt ggt atc atc act tgc cag aag cca ctc taagtctaaa 1116  
 Val Ile Lys Phe Gly Ile Ile Thr Cys Gln Lys Pro Leu  
 340 345

gctatactag gagattcaat aagactataa gagtagtgtc tcatgtgaaa gcatgaaatt 1176  
 ccttaaaaac gtcaatgtta agcctatgct tcgttatcttg ttttagataa gtatcatttc 1236  
 actcttgtct aaggtagttt ctataaacia taaataccat gaattagctc atgttatctg 1296  
 gtaaattctc ggaagtgatt gtcattggatt aactcaaaaa aaaaaaaaaa aaaa 1350

<210> 20  
 <211> 348  
 <212> PRT  
 <213> *Arabidopsis thaliana*

<400> 20  
 Met Lys Ala Thr Leu Ala Ala Pro Ser Ser Leu Thr Ser Leu Pro Tyr  
 1 5 10 15  
 Arg Thr Asn Ser Ser Phe Gly Ser Lys Ser Ser Leu Leu Phe Arg Ser  
 20 25 30  
 Pro Ser Ser Ser Ser Ser Val Ser Met Thr Thr Thr Arg Gly Asn Val  
 35 40 45  
 Ala Val Ala Ala Ala Ala Thr Ser Thr Glu Ala Leu Arg Lys Gly Ile  
 50 55 60  
 Ala Glu Phe Tyr Asn Glu Thr Ser Gly Leu Trp Glu Glu Ile Trp Gly  
 65 70 75 80  
 Asp His Met His His Gly Phe Tyr Asp Pro Asp Ser Ser Val Gln Leu  
 85 90 95  
 Ser Asp Ser Gly His Lys Glu Ala Gln Ile Arg Met Ile Glu Glu Ser  
 100 105 110

## 32

Leu Arg Phe Ala Gly Val Thr Asp Glu Glu Glu Glu Lys Lys Ile Lys  
 115 120 125  
 Lys Val Val Asp Val Gly Cys Gly Ile Gly Gly Ser Ser Arg Tyr Leu  
 130 135 140  
 Ala Ser Lys Phe Gly Ala Glu Cys Ile Gly Ile Thr Leu Ser Pro Val  
 145 150 155 160  
 Gln Ala Lys Arg Ala Asn Asp Leu Ala Ala Ala Gln Ser Leu Ser His  
 165 170 175  
 Lys Ala Ser Phe Gln Val Ala Asp Ala Leu Asp Gln Pro Phe Glu Asp  
 180 185 190  
 Gly Lys Phe Asp Leu Val Trp Ser Met Glu Ser Gly Glu His Met Pro  
 195 200 205  
 Asp Lys Ala Lys Phe Val Lys Glu Leu Val Arg Val Ala Ala Pro Gly  
 210 215 220  
 Gly Arg Ile Ile Ile Val Thr Trp Cys His Arg Asn Leu Ser Ala Gly  
 225 230 235 240  
 Glu Glu Ala Leu Gln Pro Trp Glu Gln Asn Ile Leu Asp Lys Ile Cys  
 245 250 255  
 Lys Thr Phe Tyr Leu Pro Ala Trp Cys Ser Thr Asp Asp Tyr Val Asn  
 260 265 270  
 Leu Leu Gln Ser His Ser Leu Gln Asp Ile Lys Cys Ala Asp Trp Ser  
 275 280 285  
 Glu Asn Val Ala Pro Phe Trp Pro Ala Val Ile Arg Thr Ala Leu Thr  
 290 295 300  
 Trp Lys Gly Leu Val Ser Leu Leu Arg Ser Gly Met Lys Ser Ile Lys  
 305 310 315 320  
 Gly Ala Leu Thr Met Pro Leu Met Ile Glu Gly Tyr Lys Lys Gly Val  
 325 330 335  
 Ile Lys Phe Gly Ile Ile Thr Cys Gln Lys Pro Leu  
 340 345

&lt;210&gt; 21

&lt;211&gt; 957

&lt;212&gt; DNA

&lt;213&gt; Synechocystis PCC6803

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(954)

 <223> cDNA coding for 2-methyl-6-phytylhydroquinone  
 methyltransferase



33

&lt;400&gt; 21

atg ccc gag tat ttg ctt ctg ccc gct ggc cta att tcc ctc tcc ctg	48
Met Pro Glu Tyr Leu Leu Leu Pro Ala Gly Leu Ile Ser Leu Ser Leu	
1 5 10 15	
gcg atc gcc gct gga ctg tat ctc cta act gcc cgg ggc tat cag tca	96
Ala Ile Ala Ala Gly Leu Tyr Leu Leu Thr Ala Arg Gly Tyr Gln Ser	
20 25 30	
tcg gat tcc gtg gcc aac gcc tac gac caa tgg aca gag gac ggc att	144
Ser Asp Ser Val Ala Asn Ala Tyr Asp Gln Trp Thr Glu Asp Gly Ile	
35 40 45	
ttg gaa tat tac tgg ggc gac cat atc cac ctc ggc cat tat ggc gat	192
Leu Glu Tyr Tyr Trp Gly Asp His Ile His Leu Gly His Tyr Gly Asp	
50 55 60	
ccg cca gtg gcc aag gat ttc atc caa tcg aaa att gat ttt gtc cat	240
Pro Pro Val Ala Lys Asp Phe Ile Gln Ser Lys Ile Asp Phe Val His	
65 70 75 80	
gcc atg gcc cag tgg ggc gga tta gat aca ctt ccc ccc ggc aca acg	288
Ala Met Ala Gln Trp Gly Gly Leu Asp Thr Leu Pro Pro Gly Thr Thr	
85 90 95	
gta ttg gat gtg ggt tgc ggc att ggc ggt agc agt cgc att ctc gcc	336
Val Leu Asp Val Gly Cys Gly Ile Gly Gly Ser Ser Arg Ile Leu Ala	
100 105 110	
aaa gat tat ggt ttt aac gtt acc ggc atc acc att agt ccc caa cag	384
Lys Asp Tyr Gly Phe Asn Val Thr Gly Ile Thr Ile Ser Pro Gln Gln	
115 120 125	
gtg aaa cgg gcg acg gaa tta act cct ccc gat gtg acg gcc aag ttt	432
Val Lys Arg Ala Thr Glu Leu Thr Pro Pro Asp Val Thr Ala Lys Phe	
130 135 140	
gcg gtg gac gat gct atg gct ttg tct ttt cct gac ggt agt ttc gac	480
Ala Val Asp Asp Ala Met Ala Leu Ser Phe Pro Asp Gly Ser Phe Asp	
145 150 155 160	
gta gtt tgg tcg gtg gaa gca ggg ccc cac atg cct gac aaa gct gtg	528
Val Val Trp Ser Val Glu Ala Gly Pro His Met Pro Asp Lys Ala Val	
165 170 175	
ttt gcc aag gaa tta ctg cgg gtc gtg aaa cca ggg ggc att ctg gtg	576
Phe Ala Lys Glu Leu Leu Arg Val Val Lys Pro Gly Gly Ile Leu Val	
180 185 190	
gtg gcg gat tgg aat caa cgg gac gat cgc caa gtg ccc ctc aac ttc	624
Val Ala Asp Trp Asn Gln Arg Asp Asp Arg Gln Val Pro Leu Asn Phe	
195 200 205	

## 34

tgg gaa aaa cca gtg atg cga caa ctg ttg gat caa tgg tcc cac cct 672  
 Trp Glu Lys Pro Val Met Arg Gln Leu Leu Asp Gln Trp Ser His Pro  
 210 215 220

gcc ttt gcc agc att gaa ggt ttt gcg gaa aat ttg gaa gcc acg ggt 720  
 Ala Phe Ala Ser Ile Glu Gly Phe Ala Glu Asn Leu Glu Ala Thr Gly  
 225 230 235 240

ttg gtg gag ggc cag gtg act act gct gat tgg act gta ccg acc ctc 768  
 Leu Val Glu Gly Gln Val Thr Thr Ala Asp Trp Thr Val Pro Thr Leu  
 245 250 255

ccc gct tgg ttg gat acc att tgg cag ggc att atc cgg ccc cag ggc 816  
 Pro Ala Trp Leu Asp Thr Ile Trp Gln Gly Ile Ile Arg Pro Gln Gly  
 260 265 270

tgg tta caa tac ggc att cgt ggg ttt atc aaa tcc gtg cgg gaa gta 864  
 Trp Leu Gln Tyr Gly Ile Arg Gly Phe Ile Lys Ser Val Arg Glu Val  
 275 280 285

ccg act att tta ttg atg cgc ctt gcc ttt ggg gta gga ctt tgt cgc 912  
 Pro Thr Ile Leu Leu Met Arg Leu Ala Phe Gly Val Gly Leu Cys Arg  
 290 295 300

ttc ggt atg ttc aaa gca gtg cga aaa aac gcc act caa gct taa 957  
 Phe Gly Met Phe Lys Ala Val Arg Lys Asn Ala Thr Gln Ala  
 305 310 315

<210> 22  
 <211> 318  
 <212> PRT  
 <213> Synechocystis PCC6803

<400> 22  
 Met Pro Glu Tyr Leu Leu Leu Pro Ala Gly Leu Ile Ser Leu Ser Leu  
 1 5 10 15

Ala Ile Ala Ala Gly Leu Tyr Leu Leu Thr Ala Arg Gly Tyr Gln Ser  
 20 25 30

Ser Asp Ser Val Ala Asn Ala Tyr Asp Gln Trp Thr Glu Asp Gly Ile  
 35 40 45

Leu Glu Tyr Tyr Trp Gly Asp His Ile His Leu Gly His Tyr Gly Asp  
 50 55 60

Pro Pro Val Ala Lys Asp Phe Ile Gln Ser Lys Ile Asp Phe Val His  
 65 70 75 80

Ala Met Ala Gln Trp Gly Gly Leu Asp Thr Leu Pro Pro Gly Thr Thr  
 85 90 95

Val Leu Asp Val Gly Cys Gly Ile Gly Gly Ser Ser Arg Ile Leu Ala  
 100 105 110

## 35

Lys Asp Tyr Gly Phe Asn Val Thr Gly Ile Thr Ile Ser Pro Gln Gln  
 115 120 125  
 Val Lys Arg Ala Thr Glu Leu Thr Pro Pro Asp Val Thr Ala Lys Phe  
 130 135 140  
 Ala Val Asp Asp Ala Met Ala Leu Ser Phe Pro Asp Gly Ser Phe Asp  
 145 150 155 160  
 Val Val Trp Ser Val Glu Ala Gly Pro His Met Pro Asp Lys Ala Val  
 165 170 175  
 Phe Ala Lys Glu Leu Leu Arg Val Val Lys Pro Gly Gly Ile Leu Val  
 180 185 190  
 Val Ala Asp Trp Asn Gln Arg Asp Asp Arg Gln Val Pro Leu Asn Phe  
 195 200 205  
 Trp Glu Lys Pro Val Met Arg Gln Leu Leu Asp Gln Trp Ser His Pro  
 210 215 220  
 Ala Phe Ala Ser Ile Glu Gly Phe Ala Glu Asn Leu Glu Ala Thr Gly  
 225 230 235 240  
 Leu Val Glu Gly Gln Val Thr Thr Ala Asp Trp Thr Val Pro Thr Leu  
 245 250 255  
 Pro Ala Trp Leu Asp Thr Ile Trp Gln Gly Ile Ile Arg Pro Gln Gly  
 260 265 270  
 Trp Leu Gln Tyr Gly Ile Arg Gly Phe Ile Lys Ser Val Arg Glu Val  
 275 280 285  
 Pro Thr Ile Leu Leu Met Arg Leu Ala Phe Gly Val Gly Leu Cys Arg  
 290 295 300  
 Phe Gly Met Phe Lys Ala Val Arg Lys Asn Ala Thr Gln Ala  
 305 310 315

&lt;210&gt; 23

&lt;211&gt; 1395

&lt;212&gt; DNA

&lt;213&gt; Nicotiana tabacum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1392)

 <223> cDNA coding for geranylgeranylpyrophosphate  
 oxidoreductase

&lt;400&gt; 23

atg gct tcc att gct ctc aaa act ttc acc ggc ctc cgt caa tcc tcg 48  
 Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser Ser  
 1 5 10 15

## 36

ccg gaa aac aat tcc att act ctt tct aaa tcc ctc ccc ttc acc caa	96
Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe Thr Gln	
20 25 30	
acc cac cgt agg ctc cga atc aat gct tcc aaa tcc agc cca aga gtc	144
Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser Pro Arg Val	
35 40 45	
aac ggc cgc aat ctt cgt gtt gcg gtg gtg ggc ggt ggt cct gct ggt	192
Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly Gly Pro Ala Gly	
50 55 60	
ggc gcc gcc gct gaa aca ctc gcc aag gga gga att gaa acc ttc tta	240
Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly Ile Glu Thr Phe Leu	
65 70 75 80	
atc gaa cgc aaa atg gac aac tgc aaa ccc tgc ggt ggg gcc atc cca	288
Ile Glu Arg Lys Met Asp Asn Cys Lys Pro Cys Gly Gly Ala Ile Pro	
85 90 95	
ctt tgc atg gtg gga gaa ttt gac ctc cct ttg gat atc att gac cgg	336
Leu Cys Met Val Gly Glu Phe Asp Leu Pro Leu Asp Ile Ile Asp Arg	
100 105 110	
aaa gtt aca aag atg aag atg att tcc cca tcc aac gtt gct gtt gat	384
Lys Val Thr Lys Met Lys Met Ile Ser Pro Ser Asn Val Ala Val Asp	
115 120 125	
att ggt cag act tta aag cct cac gag tac atc ggt atg gtg cgc cgc	432
Ile Gly Gln Thr Leu Lys Pro His Glu Tyr Ile Gly Met Val Arg Arg	
130 135 140	
gaa gta ctc gat gct tac ctc cgt gac cgc gct gct gaa gcc gga gcc	480
Glu Val Leu Asp Ala Tyr Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala	
145 150 155 160	
tct gtt ctc aac ggc ttg ttc ctc aaa atg gac atg ccc aaa gct ccc	528
Ser Val Leu Asn Gly Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro	
165 170 175	
aac gca cct tac gtc ctt cac tac aca gct tac gac tcc aaa act aat	576
Asn Ala Pro Tyr Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn	
180 185 190	
ggc gcg ggg gag aag cgt acc ctg gaa gtt gac gcc gtt atc ggc gct	624
Gly Ala Gly Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala	
195 200 205	
gac ggt gca aat tcc cgt gtc gca aaa tcc ata aac gcc ggt gac tac	672
Asp Gly Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr	
210 215 220	

gag	tac	gct	att	gca	ttc	caa	gaa	agg	att	aaa	att	tcc	gat	gat	aaa	720
Glu	Tyr	Ala	Ile	Ala	Phe	Gln	Glu	Arg	Ile	Lys	Ile	Ser	Asp	Asp	Lys	
225					230					235					240	
atg	aag	tat	tac	gag	aat	tta	gct	gaa	atg	tac	gtg	ggt	gat	gac	gtg	768
Met	Lys	Tyr	Tyr	Glu	Asn	Leu	Ala	Glu	Met	Tyr	Val	Gly	Asp	Asp	Val	
				245					250					255		
tcc	cct	gat	ttt	tac	ggg	tgg	gtt	ttc	ccc	aaa	tgt	gac	cac	gtt	gcc	816
Ser	Pro	Asp	Phe	Tyr	Gly	Trp	Val	Phe	Pro	Lys	Cys	Asp	His	Val	Ala	
			260					265					270			
gtt	ggc	act	ggc	aca	gtc	acc	cac	aaa	gct	gac	atc	aaa	aaa	ttc	cag	864
Val	Gly	Thr	Gly	Thr	Val	Thr	His	Lys	Ala	Asp	Ile	Lys	Lys	Phe	Gln	
		275					280					285				
cta	gct	aca	aga	ttg	aga	gct	gat	tcc	aaa	atc	acc	ggc	gga	aaa	att	912
Leu	Ala	Thr	Arg	Leu	Arg	Ala	Asp	Ser	Lys	Ile	Thr	Gly	Gly	Lys	Ile	
	290					295					300					
atc	cgg	gtc	gag	gcc	cac	ccg	att	cca	gaa	cac	cca	aga	ccc	aga	aga	960
Ile	Arg	Val	Glu	Ala	His	Pro	Ile	Pro	Glu	His	Pro	Arg	Pro	Arg	Arg	
305				310						315				320		
tta	caa	gac	aga	gtt	gca	ttg	gtt	ggt	gat	gcg	gca	ggg	tac	gtg	acc	1008
Leu	Gln	Asp	Arg	Val	Ala	Leu	Val	Gly	Asp	Ala	Ala	Gly	Tyr	Val	Thr	
			325					330					335			
aaa	tgt	tcg	ggc	gaa	ggg	att	tac	ttc	gcg	gca	aag	agt	gga	cgt	atg	1056
Lys	Cys	Ser	Gly	Glu	Gly	Ile	Tyr	Phe	Ala	Ala	Lys	Ser	Gly	Arg	Met	
		340						345				350				
tgt	gct	gaa	gca	att	gtt	gaa	ggg	tca	gaa	atg	gga	aaa	aga	atg	gtg	1104
Cys	Ala	Glu	Ala	Ile	Val	Glu	Gly	Ser	Glu	Met	Gly	Lys	Arg	Met	Val	
	355					360					365					
gac	gag	agt	gat	ttg	agg	aag	tat	ttg	gag	aaa	tgg	gac	aag	act	tat	1152
Asp	Glu	Ser	Asp	Leu	Arg	Lys	Tyr	Leu	Glu	Lys	Trp	Asp	Lys	Thr	Tyr	
	370				375					380						
tgg	cca	acg	tac	aag	gtg	ctt	gat	ata	ttg	cag	aag	gta	ttt	tac	agg	1200
Trp	Pro	Thr	Tyr	Lys	Val	Leu	Asp	Ile	Leu	Gln	Lys	Val	Phe	Tyr	Arg	
385				390						395			400			
tcg	aat	ccg	gcg	agg	gaa	gca	ttt	gtt	gaa	atg	tgc	gca	gat	gag	tat	1248
Ser	Asn	Pro	Ala	Arg	Glu	Ala	Phe	Val	Glu	Met	Cys	Ala	Asp	Glu	Tyr	
		405						410				415				
gtg	cag	aag	atg	aca	ttt	gac	agc	tat	ttg	tac	aag	aaa	gta	gca	cca	1296
Val	Gln	Lys	Met	Thr	Phe	Asp	Ser	Tyr	Leu	Tyr	Lys	Lys	Val	Ala	Pro	

gga aac cca att gaa gac ttg aag ctt gct gtg aat acc att gga agt 1344  
Gly Asn Pro Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser  
435 440 445

ttg gtg aga gct aat gca cta aga agg gaa atg gac aag ctc agt gta 1392  
Leu Val Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val  
450 455 460

taa 1395

<210> 24

<211> 464

<212> PRT

<213> Nicotiana tabacum

<400> 24

Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser Ser  
1 5 10 15

Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe Thr Gln  
20 25 30

Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser Pro Arg Val  
35 40 45

Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly Gly Pro Ala Gly  
50 55 60

Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly Ile Glu Thr Phe Leu  
65 70 75 80

Ile Glu Arg Lys Met Asp Asn Cys Lys Pro Cys Gly Gly Ala Ile Pro  
85 90 95

Leu Cys Met Val Gly Glu Phe Asp Leu Pro Leu Asp Ile Ile Asp Arg  
100 105 110

Lys Val Thr Lys Met Lys Met Ile Ser Pro Ser Asn Val Ala Val Asp  
115 120 125

Ile Gly Gln Thr Leu Lys Pro His Glu Tyr Ile Gly Met Val Arg Arg  
130 135 140

Glu Val Leu Asp Ala Tyr Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala  
145                      150                      155                      160

Ser Val Leu Asn Gly Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro  
165 170 175

Asn Ala Pro Tyr Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn  
180 185 190

Gly Ala Gly Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala  
195 200 205

## 39

Asp Gly Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr  
 210 215 220  
 Glu Tyr Ala Ile Ala Phe Gln Glu Arg Ile Lys Ile Ser Asp Asp Lys  
 225 230 235 240  
 Met Lys Tyr Tyr Glu Asn Leu Ala Glu Met Tyr Val Gly Asp Asp Val  
 245 250 255  
 Ser Pro Asp Phe Tyr Gly Trp Val Phe Pro Lys Cys Asp His Val Ala  
 260 265 270  
 Val Gly Thr Gly Thr Val Thr His Lys Ala Asp Ile Lys Lys Phe Gln  
 275 280 285  
 Leu Ala Thr Arg Leu Arg Ala Asp Ser Lys Ile Thr Gly Gly Lys Ile  
 290 295 300  
 Ile Arg Val Glu Ala His Pro Ile Pro Glu His Pro Arg Pro Arg Arg  
 305 310 315 320  
 Leu Gln Asp Arg Val Ala Leu Val Gly Asp Ala Ala Gly Tyr Val Thr  
 325 330 335  
 Lys Cys Ser Gly Glu Gly Ile Tyr Phe Ala Ala Lys Ser Gly Arg Met  
 340 345 350  
 Cys Ala Glu Ala Ile Val Glu Gly Ser Glu Met Gly Lys Arg Met Val  
 355 360 365  
 Asp Glu Ser Asp Leu Arg Lys Tyr Leu Glu Lys Trp Asp Lys Thr Tyr  
 370 375 380  
 Trp Pro Thr Tyr Lys Val Leu Asp Ile Leu Gln Lys Val Phe Tyr Arg  
 385 390 395 400  
 Ser Asn Pro Ala Arg Glu Ala Phe Val Glu Met Cys Ala Asp Glu Tyr  
 405 410 415  
 Val Gln Lys Met Thr Phe Asp Ser Tyr Leu Tyr Lys Lys Val Ala Pro  
 420 425 430  
 Gly Asn Pro Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser  
 435 440 445  
 Leu Val Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val  
 450 455 460

&lt;210&gt; 25

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; oligonucleotide

<400> 25	
gtcgcacggnc cnatnggngc naangg	26
<210> 26	
<211> 24	
<212> DNA	
<213> oligonucleotide	
<400> 26	
aagcttccga tctagtaaca taga	24
<210> 27	
<211> 32	
<212> DNA	
<213> oligonucleotide	
<400> 27	
attctagaca tggagtcaaa gattcaaata ga	32
<210> 28	
<211> 32	
<212> DNA	
<213> oligonucleotide	
<400> 28	
attctagagg acaatcagta aattgaacgg ag	32
<210> 29	
<211> 26	
<212> DNA	
<213> oligonucleotide	
<400> 29	
atgtcgacat gtcttatgtt accgat	26
<210> 30	
<211> 25	
<212> DNA	
<213> oligonucleotide	
<400> 30	
atggatccct ggttcatatg ataca	25
<210> 31	
<211> 26	
<212> DNA	
<213> oligonucleotide	
<400> 31	
atgtcgacgg aaactctgaa ccatat	26



<210> 32  
<211> 25  
<212> DNA  
<213> oligonucleotide

<400> 32  
atggtaccga atgtgatgcc taagt 25

<210> 33  
<211> 29  
<212> DNA  
<213> oligonucleotide

<400> 33  
ggtacctcra acatraangc catngtncc 29

<210> 34  
<211> 25  
<212> DNA  
<213> oligonucleotide

<400> 34  
gaattcgatc tgcgtctca aactc 25

<210> 35  
<211> 26  
<212> DNA  
<213> oligonucleotide

<400> 35  
ggtaccgtga tagtaaaca ctaatg 26

<210> 36  
<211> 34  
<212> DNA  
<213> oligonucleotide

<400> 36  
atggtacctt ttttgataa acttatcttc atag 34

<210> 37  
<211> 43  
<212> DNA  
<213> oligonucleotide

<400> 37  
atgtcgaccc gggatccagg gccctgatgg gtcccatttt ccc 43

<210> 38  
<211> 25  
<212> DNA  
<213> oligonucleotide

<400> 38  
 gtcgacgaat ttccccgaat cgttc 25

<210> 39  
 <211> 24  
 <212> DNA  
 <213> oligonucleotide

<400> 39  
 aagcttccga tctagtaaca taga 24

<210> 40  
 <211> 25  
 <212> DNA  
 <213> oligonucleotide

<400> 40  
 aagcttgatc tgtcgtctca aactc 25

<210> 41  
 <211> 1721  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> misc\_feature  
 <222> (1)..(8)  
 <223> restriction site linker

<220>  
 <221> misc\_feature  
 <222> (1714)..(1721)  
 <223> restriction site linker

<220>  
 <221> misc\_feature  
 <222> (9)..(1713)  
 <223> fragment from gene coding for maleylacetoactetate  
 isomerase (MAAI)

<400> 41  
 atgtcgacat gtcttatgtt accgattttt atcagggcgaa gttgaagctc tactcttact 60  
 ggagaagctc atgtgctcat cgcgtccgta tcgcctcac tttaaagggt accagccaat 120  
 gattttattc ttttcttgag agcaattctt tgatctgaat ttggttcttg ttcgattttc 180  
 attagggctt gattatgaat atataccggt taatttgctc aaaggggatc aatccgattc 240  
 aggtgcgtag tttctaggtt atattgaact ttatttgaag taacattgta aagataagaa 300  
 tggtaagtaa ctgagatttc ttatgttaga cttagaagtt tattcgtttt gggtctctag 360  
 atttcaagaa gatcaatcca atgggcactg taccagcgct tgttgatggg gatgttgtga 420  
 ttaatgactc tttcgcaata ataatgggtc gtagtaacac atccatttag tttgtttggg 480  
 tttgttgatg aaaaggaaca ttcgtttatt cgtcttggtt tttttcaa atggacagtacc 540  
 tggatgataa gtatccggag ccaccgctgt taccagtgta ctaccataaa cgggcggtaa 600  
 attaccaggt atcttcgatc ctttgtcttc agatgatgat gtgttgccat catctgcaaa 660  
 accatgtagt taagtccaaa tgtagtgaac attatcagct ttagatttgc agtgtgatcg 720

## 43

```

ttgttcttat tttgtatatt tcaggcgacg agtattgtca tgtctggtat acagcctcat 780
caaaatatgg ctctttttgt gagaagatga gattaatgta atggattcta ctaatggagg 840
ttctataaca aagcaaacat agttacattt tgtcattttt tttaacagag gtatctcgag 900
gacaagataa atgctgagga gaaaactgct tggattacta atgctatcac aaaaggattc 960
acaggatga tatctctaata ctacctatac gtaatcaaga accaagacat atgttcaaaa 1020
tgtgattttg ttgatattgt gggtgtacag gtttataacg acctgtctga taatgtctca 1080
tatgtccttc agctctcgag aaactgttgg tgagttgctg tggaaaatac gcgactggtg 1140
atgaagttaa cttgggtatgt ctctaaatct ccctggataa tctctatggt actactctct 1200
tctttattac aatgaagcat tgttttgacg gctgactctt tcctagcacc acagatccac 1260
gcagcattca acagattcca tattaacatg gtacttttcc tcagctaate tcttctctcg 1320
gtacctagat attgcattgt atatcccccc aaattccatg gaatccttga tcagagtttt 1380
aaggtagcat gaaccaaagc ttatctctgt ctacacttt cacattcaca gagtaacata 1440
gacgtaatac tcagtttcat aacttttttt cctcgcatca cttgggtttc atctctacaa 1500
ttttgttgta taggaacat tcccgcactc tgcaagggtt tacgagtcac acaacgaact 1560
gcctgcattt caaatgcag tcccggagaa gcaaccagat actccttcca ccatctgatt 1620
ctgtgaaccg taagcttctc tcagtctcag ctcaataaaa tctcttagga aacaacaaca 1680
acaccttgaa cttaaagtga tcatatgaac cagggatcca t 1721

```

&lt;210&gt; 42

&lt;211&gt; 622

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)..(8)

&lt;223&gt; restriction site linker

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (615)..(622)

&lt;223&gt; restriction site linker

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (9)..(614)

<223> fragment from gene coding for fumarylacetoacetate  
hydrolase (FAAH)

&lt;400&gt; 42

```

atgtcgacgg aaactctgaa ccatattttg gaccttcgaa gaaacttgat tttgagcttg 60
agatggtaag catctgatgc ctgagttatg tggatttggt ttacaatgat tcggttgatg 120
ctttttggtg ctagttaaga ataacggcat tgacaaacct ctcttttatc acatgatatt 180
caggctgctg tgggttggtc aggaaatgaa ttgggaaagc ctattgacgt gaataatgca 240
gccgatcata tatttggtct attactgatg aatgactgga gtggtactca cttaactata 300
gttttcggtg agtcatcttt aacctgaccg ggcatgaccg gtttttttaa atgtttggtg 360
ttatagctag ggatattcag gcgtgggagt atgtacctct tgggtccttc ctggggaaga 420
gttttggtga gatatttgcc ttcaatactt tgatttcatt tcctctagtt gaagtatatg 480
ggcaaagaac ttcggtgaat gttgtcttgt tgtgtgttag ggactactat atccccctgg 540
attgttacct tggatgcgct tgagcctttt gggtgtcaag ctccaagca ggttggtact 600
taggcacac attcggtacc at 622

```

<210> 43

<211> 32

<212> DNA

<213> oligonucleotide

<400> 43

atgaattcca tggagtcaaa gattcaaata ga

32

<210> 44

<211> 32

<212> DNA

<213> oligonucleotide

<400> 44

atgaattcgg acaatcagta aattgaacgg ag

32